The EMBO Journal vol.14 no.17 pp.4365–4373, 1995

The products of the *SUP45* (eRF1) and *SUP35* genes interact to mediate translation termination in *Saccharomyces cerevisiae*

Ian Stansfield, Kerrie M.Jones, Vitaly V.Kushnirov[1], Adilya R.Dagkesamanskaya[1], Andrey I.Poznyakovski[1], Sergey V.Paushkin[1], Concepcion R.Nierras[2,3], Brian S.Cox[2,4] Michael D.Ter-Avanesyan[1] and Mick F.Tuite[5]

Research School of Biosciences, University of Kent, Canterbury CT2 7NJ, UK, [1]Institute of Experimental Cardiology, 3rd Cherepkovskaya Street 15A, 121552, Moscow, Russia and [2]Micro Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK
[3]Present address: Department of Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY, USA
[4]Present address: University of Oxford, Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DU, UK

[5]Corresponding author

The product of the yeast *SUP45* gene (Sup45p) is highly homologous to the *Xenopus* eukaryote release factor 1 (eRF1), which has release factor activity *in vitro*. We show, using the two-hybrid system, that in *Saccharomyces cerevisiae* Sup45p and the product of the *SUP35* gene (Sup35p) interact *in vivo*. The ability of Sup45p C-terminally tagged with $(His)_6$ to specifically precipitate Sup35p from a cell lysate was used to confirm this interaction *in vitro*. Although overexpression of either the *SUP45* or *SUP35* genes alone did not reduce the efficiency of codon-specific tRNA nonsense suppression, the simultaneous overexpression of both the *SUP35* and *SUP45* genes in nonsense suppressor tRNA-containing strains produced an antisuppressor phenotype. These data are consistent with Sup35p and Sup45p forming a complex with release factor properties. Furthermore, overexpression of either *Xenopus* or human eRF1 (*SUP45*) genes also resulted in antisuppression only if that strain was also overexpressing the yeast *SUP35* gene. Antisuppression is a characteristic phenotype associated with overexpression of both prokaryote and mitochondrial release factors. We propose that Sup45p and Sup35p interact to form a release factor complex in yeast and that Sup35p, which has GTP binding sequence motifs in its C-terminal domain, provides the GTP hydrolytic activity which is a demonstrated requirement of the eukaryote translation termination reaction.

Keywords: release factor/*Saccharomyces cerevisiae*/*SUP35*/*SUP45*/translation termination

Introduction

In mRNAs the three codons UAA, UAG and UGA are almost universally employed to signal termination of translation. In bacteria this process is catalysed by one of two release factors, RF1 at UAA and UAG codons and RF2 at UAA and UGA (Scolnick *et al.*, 1968). A third release factor, RF3, showing some homology to bacterial elongation factor EF-G, enhances the rate of RF1 and RF2 catalysed termination in a GTP-dependent and codon-independent manner (Milman *et al.*, 1969; Grentzmann *et al.*, 1994; Mikuni *et al.*, 1994). Following peptidyl-tRNA hydrolysis to free the nascent peptide, a ribosome release factor (RRF) is employed to release the ribosomal subunits from the mRNA, enabling them to participate in new rounds of initiation (Hirashima and Kaji, 1972).

In eukaryotes the process of translational termination has been defined less equivocally: it is known that a single polypeptide, the eukaryote release factor (eRF) will catalyse termination *in vitro* at all three stop codons in a GTP-dependent manner (Goldstein *et al.*, 1970; Konecki *et al.*, 1977). An additional stimulatory factor 's' was identified, but not further characterized (Konecki *et al.*, 1977). Attempts to clone the gene coding for eRF resulted in isolation of a gene with 89% amino acid identity to bovine tryptophanyl-tRNA synthetase (TrpRS; Lee *et al.*, 1990). However, it was subsequently shown that the TrpRS polypeptide purified from rabbit reticulocyte lysates does not have eRF activity, but rather co-purifies with a protein that does (Frolova *et al.*, 1993). Using an *in vitro* peptidyl release assay, originally employed to identify components of the bacterial termination complex (Caskey *et al.*, 1968, 1973), Frolova *et al.*, (1994) showed that the eRF protein had release factor activity in response to all three termination codons. N-Terminal sequencing identified the protein as a product of the *SUP45* gene family coding for closely-related polypeptides, including the *Saccharomyces cerevisiae* Sup45p (Himmelfarb *et al.*, 1985), *Xenopus* C11 (Tassan *et al.*, 1993) and human TB3-1 (Grenett *et al.*, 1992) proteins. Frolova *et al.* (1994) showed that these latter two proteins also have omnipotent release factor activity.

The yeast *SUP45* gene codes for a protein (Sup45p) that is tightly associated with polysomal ribosomes (Stansfield *et al.*, 1992) and present at a low level in the cell, with a molar ratio to ribosomes of $<1:20$ (Stansfield *et al.*, 1992), an abundance typical of the *Escherichia coli* release factors RF1 and RF2 (Klein and Capecchi, 1971). Mutant alleles of the *SUP45* gene which exhibit either omnipotent suppressor (*sup1*, Inge-Vechtomov and Andrianova, 1970; *sup45*, Hawthorne and Leupold, 1974) or allosuppressor phenotypes (*sal4*, Cox, 1977) are known. Allosuppressors are selected on the basis of an ability to enhance the suppressor efficiency of the weak ochre suppressor tRNA *SUQ5*, while omnipotent suppression is thought to result when the limited nonsense suppressor ability of some wild-type tRNAs is enhanced (Stansfield *et al.*, 1995). Both the omnipotent and allosuppressor phenotypes can

be associated with a single *sup45* allele (Stansfield *et al.*, 1995), as would be expected of a gene encoding a protein with release factor activity.

Although the release factor activity of yeast Sup45p protein has not been formally demonstrated using the *in vitro* biochemical termination assay, it seems extremely likely that it does perform this role *in vivo*, for two reasons: first, the yeast Sup45p polypeptide exhibits a high degree of amino acid identity with the *Xenopus* (C11) and human (TB-3) eRF1 proteins (68 and 66% respectively; Frolova *et al.*, 1994); second, the *Xenopus*, human and Syrian hamster eRF1 genes can be used to functionally replace the *S.cerevisiae SUP45* gene *in vivo* (B.Urbero, L.Eurwilaichitr, I.Stansfield, M.Philippe, M.Kress and M.F.Tuite, in preparation).

A eukaryote gene family has thus been identified which codes for proteins with omnipotent release factor activity. However, the translation termination process is known to be GTP dependent in rabbit reticulocyte systems (Goldstein *et al.*, 1970; Konecki *et al.*, 1977), yet the *SUP45* gene family members show no homology to any known GTP binding sequence motifs (I.Stansfield and M.F.Tuite, unpublished data). This raises the question of whether eRF1 (Sup45p) is the only component of the eukaryotic release factor. One potential candidate protein with GTP binding sequence motifs, which by virtue of its mutant phenotypes may play a role in translation termination, is the product of the *SUP35* gene (Sup35p; Tuite and Stansfield, 1994). Like the *SUP45* gene, the *SUP35* gene was identified through both allosuppressor and omnipotent suppressor mutational screens and, like Sup45p, Sup35p is also closely associated with the ribosome at an approximate stoichiometry of 1 mol Sup35p to 20 mol ribosomes (Didichenko *et al.*, 1991). In the following work we confirm that Sup35p and Sup45p interact both *in vivo* and *in vitro* and that they act in concert to form a functional termination complex *in vivo* in *S.cerevisiae*.

Results

Sup35p and Sup45p can interact in vivo

The yeast *SUP35* and *SUP45* genes have a number of properties in common. Both genes were identified through screens for omnipotent suppressor mutations (Inge-Vechtomov and Andrianova, 1970; Hawthorne and Leupold, 1974) and for allosuppressor mutations (Cox, 1977) and both gene products Sup35p and Sup45p are associated with polysomal ribosomes, with a probable location on the 40S subunit (Didichenko *et al.*, 1991; Stansfield *et al.*, 1992). Two observations indicate that these two proteins either participate in a common process or interact directly: first, a *sal3 sal4* (*sup35 sup45*) double allosuppressor mutant is inviable (Cox, 1977); second, an increased dosage of the *SUP45* gene relieves the degree of temperature sensitivity of some *sup35* conditional-lethal mutations, while overexpression of *SUP35* partially complements a *sup45* temperature-sensitive mutation (Ter-Avanesyan *et al.*, 1984, 1993).

In order to confirm that Sup35p and Sup45p interact *in vivo*, the *GAL4*-based two-hybrid system was employed as a direct assay of protein–protein interaction *in vivo* (Fields and Song, 1989; Chien *et al.*, 1991). The *SUP35* and *SUP45* genes were cloned in-frame downstream of the *GAL4* activation and binding domains respectively (see Materials and methods). The two plasmids thus generated, together with a number of control constructs detailed below, were transformed in different pair-wise combinations into the *S.cerevisiae* host strain Y526 (Bartel *et al.*, 1993a), carrying the *GAL1–lacZ* gene fusion to assay Gal4p activation of *GAL1* transcription (Fields and Song, 1989). The activity of the β-galactosidase reporter was assayed initially by colony colour, using the chromogenic substrate 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal). For each of the different classes of transformant detailed in Figure 1, five independent double transformants were screened, including a number of pair-wise control combinations of plasmids. The results showed that neither Sup45p–bd nor Sup35p–ad fusion proteins alone interacted with the complementary *GAL4* domains, the colonies of these double transformants remaining white on X-gal medium (not shown). There was also no non-specific interaction of Sup45p–bd or Sup35p–ad fusions with other proteins, exemplified by *GAL4* domain fusions with the dimerization domain of transcription factor *CPF1* (Dowell *et al.*, 1992). Only double transformants expressing both bd–Sup45p and ad–Sup35p fusions (pUKC601/pUKC605) gave a blue colony colour with X-gal (not shown).

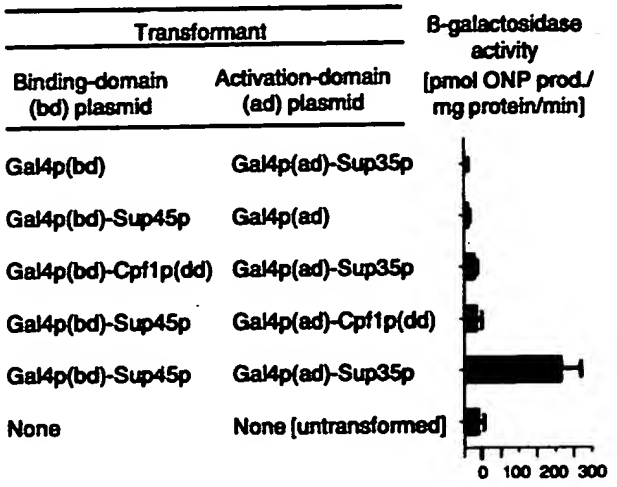


Fig. 1. Two-hybrid analysis of Sup35p interaction with Sup45p. β-Galactosidase activities of cell lysates of strain Y526 both untransformed and transformed with pair-wise combinations of the two-hybrid vectors were determined (where background activity represented by the untransformed Y526 strain was taken as 0 pmol/min/mg protein). β-Galactosidase activities are the mean of three determinations. Error bars represent ± 1 SD. Domains of proteins expressed, in terms of amino acid numberings, are given in Table II.

These results were confirmed using the β-galactosidase assay (Figure 1), with all control cell lysates exhibiting β-galactosidase activities approximately equivalent to the background levels of activity measured in the untransformed Y526 host strain lysate. However, the β-galactosidase activity of the lysate expressing the bd–Sup45p plus ad–Sup35p protein fusions was measured at 227 pmol/min/mg protein above background, strong evidence that Sup45p and Sup35p interact directly *in vivo*. Moreover, this activity was twice that measured in a positive control transformant expressing both the ad and bd fusions of

Snf1p and Snf4p respectively (not shown), two proteins whose interaction has been previously demonstrated using the two-hybrid system (Fields and Song, 1989).

Sup45p and Sup35p also interact in vitro

To confirm that Sup45p and Sup35p specifically interact, the ability of immobilized Sup45p to precipitate Sup35p from a post-ribosomal supernatant (PRS) was tested. The immobilization of Sup45p onto nickel (Ni–NTA)–agarose beads was achieved by tagging the polypeptide at its C-terminus with a $(His)_6$ peptide (see Materials and methods).

The multicopy plasmid pUKC625 encoding the Sup45p$(His)_6$ construct was transformed into yeast strain MT552/36d and a PRS prepared. Since Sup35p and Sup45p are both known to bind to the 40S ribosomal subunit and 80S ribosomes (Didichenko *et al.*, 1991; Stansfield *et al.*, 1992), the preparation of a PRS fraction by removal of ribosomal material from the Sup45p$(His)_6$ lysate was felt to be essential for assessing interactions between the two polypeptides in solution.

The PRS, containing Sup45p$(His)_6$ in high salt concentration (0.8 M KCl) lysis buffer, was repeatedly passed through Ni–NTA–agarose, ensuring preferential binding of Sup45p$(His)_6$ to the resin and limiting non-specific binding of other proteins. The majority of non-specifically bound proteins were removed by washing the resin under stringent conditions of low pH and high salt and glycerol (Materials and methods). This resulted in a preparation of partially purified Sup45p$(His)_6$ bound to the nickel resin. A control nickel resin was prepared in the same way using a PRS prepared from untransformed strain MT552/36d. This control resin had the same non-specific complement of PRS proteins bound to it, but lacked any bound Sup45p$(His)_6$ detectable by immunoblot assay.

In order to investigate the Sup45p–Sup35p interaction, the Sup45p$(His)_6$–resin and the control resin were tested for their ability to precipitate Sup35p from solution. Sup35p was overexpressed in yeast strain MT422/1c transformed with plasmid pUKC606, from which a PRS (Sup35p enriched) was prepared. The Sup35p-rich PRS was then incubated with either the Sup45p$(His)_6$–resin or the control resin for 2 h in a low salt buffer to promote protein–protein interactions. Both control resin and Sup45p$(His)_6$–resin were then returned to columns to facilitate washing.

Samples of the Sup35p-rich PRS were analysed by SDS–PAGE and Western blotting before and after incubation with the two resin preparations, as were, after extensive washing, samples of the Sup45p$(His)_6$–resin and the control resin. The Western blots were probed with antibodies to Sup35p and Sup45p (Figure 2b and c). The results show that the Sup45p$(His)_6$ was immobilized on the Ni–NTA resin, whilst no Sup45p was detectable on the control resin (Figure 2b). The small amounts of Sup45p present in the Sup35p PRS, derived from expression of the genomic copy of *SUP45*, were not detectable under the conditions used.

Sup35p was detectable in the Sup35p-rich PRS (Figure 2c) and to a lesser extent in the supernatants following incubation with the two resin types. This loss of Sup35p from the PRS was due to non-specific adsorbtion of Sup35p onto the control resin (data not shown) and specific interaction of Sup35p with Sup45p on the Sup45p$(His)_6$–resin. This difference in the nature of Sup35p interaction with the two resins was clearly demonstrated by the fact that while no Sup35p was retained on the control resin following washing, Sup35p remained tightly bound to the Sup45p$(His)_6$–resin (Figure 2c). These data confirm that Sup35p interacts tightly and specifically with Sup45p in solution. In a complementary experiment which employed the reverse strategy, a glutathione S-transferase–Sup35p fusion protein bound to glutathione–agarose beads was used to precipitate Sup45p from a cell lysate, further confirming that Sup35p and Sup45p interact *in vitro* (data not shown).

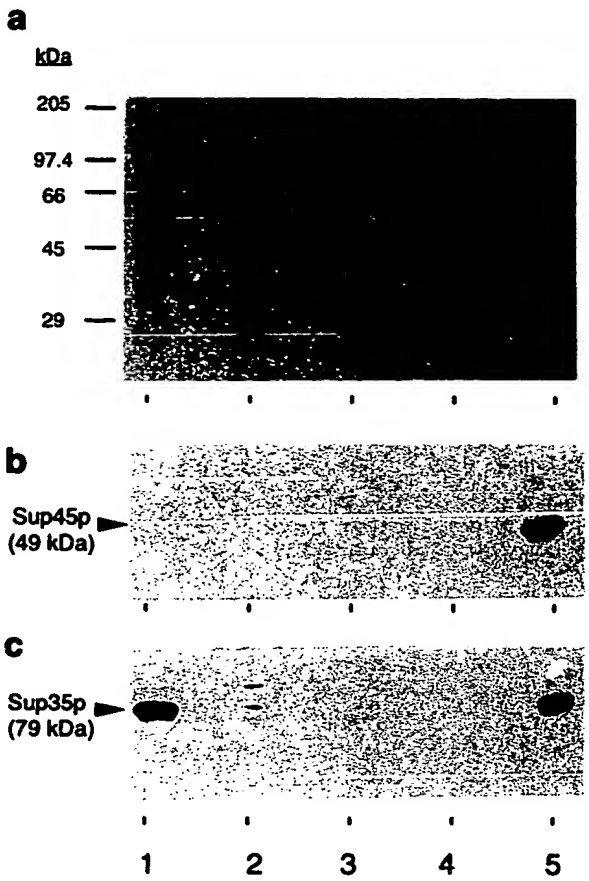


Fig. 2. Precipitation of Sup35p from a post-ribosomal supernatant (PRS) by immobilized Sup45p $(His)_6$. The PRS fraction from a yeast strain overexpressing Sup35p was incubated with either Ni–NTA resin loaded with Sup45p$(His)_6$ or a Sup45p-free control resin (Materials and methods). Following washing, equal quantities of the two resins and post-incubation supernatant fractions were assayed for total protein by Coomassie blue staining of SDS–PAGE gels (a) and the presence of Sup45p (b) or Sup35p (c) detected on immunoblots probed with Sup45p and Sup35p affinity-purified antibodies (Materials and methods). Lane 1, Sup35p-enriched PRS; lane 2, supernatant following incubation with control resin; lane 3, control resin following washing; lane 4, supernatant following incubation with Sup45p$(His)_6$–resin; lane 5, resin-bound Sup45p$(His)_6$ following washing. The positions of migration of molecular mass markers are shown, as are the positions of the bands representing Sup45p and Sup35p.

Multicopy expression phenotypes of the SUP45 and SUP35 genes

The rabbit, *Xenopus* and human homologues of yeast Sup45p have peptidyl release activity in an *in vitro* assay

Double Transformant			Relevant protein(s) expressed			Three independent double-transformants grown on defined medium (-ura -leu)	
						+adenine	-adenine
pUKC606	+	YEp24	Sup35p(Sc)		-		
pRS425	+	pUKC802	-		Sup45p(Sc)		
pUKC606	+	pUKC802	Sup35p(Sc)	+	Sup45p(Sc)		
pUKC606	+	pEO6	Sup35p(Sc)	+	eRF1(Hs)		
pUKC606	+	pEX1	Sup35p(Sc)	+	eRF1(Xl)		

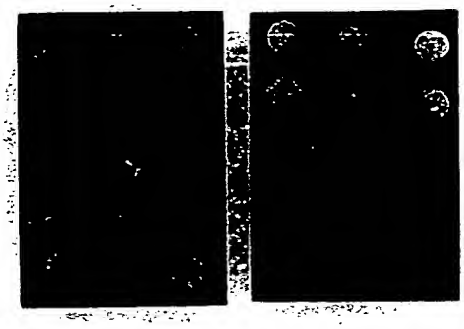


Fig. 3. Simultaneous overexpression of Sup35p and Sup45p acts to antisuppress the ochre suppressor tRNA *SUP4*. Strain MT422/1c carrying the *SUP4-o* tRNA ochre suppressor and the *ade2-1* ochre allele was transformed with the pair-wise combinations of plasmids indicated and growth of three independent transformants on the control medium (+ adenine) compared with growth on defined medium lacking adenine. Each yeast 'colony' represents ~1×10^5 cells spotted onto the defined medium in 3 μl water and allowed to grow for 3 days at 30°C. Expressed proteins coded for by *S.cerevisiae* genes are designated Sc, by *Xenopus laevis* cDNA, Xl and by human cDNA, Hs.

of the translation termination reaction (Frolova *et al.*, 1994). One characteristic typical of release factors is that their overexpression produces an antisuppressor phenotype in a *SUP* tRNA genetic background, by out-competing suppressor tRNAs during stop codon binding (Weiss,R.B. *et al.*, 1984; Pel *et al.*, 1992). It might therefore be expected that the yeast *SUP45* gene would act as an antisuppressor when overexpressed, given the reasonable assertion that it too encodes a protein with eRF1 activity (Tassan *et al.*, 1993; B.Urbero, L.Eurwilaichitr, I.Stansfield, M.Philippe, M.Kress and M.F.Tuite, in preparation). However, overexpression of the yeast *SUP45* gene from a multicopy plasmid does not generate an antisuppressor phenotype (I.Stansfield and M.F.Tuite, unpublished data) The *SUP35* gene also does not act as an antisuppressor when overexpressed, instead causing increased suppression of nonsense codons (Chernoff *et al.*, 1993). We therefore examined whether simultaneous overexpression of both *SUP35* and *SUP45* genes could produce an antisuppressor phenotype: a positive result would indicate that *in vivo* Sup35p and Sup45p form a complex with release factor properties.

To test this hypothesis we employed two multicopy plasmids to overexpress the Sup35p and Sup45p proteins simultaneously in a yeast strain carrying the efficient ochre suppressor tRNA mutation *SUP4*, testing for the antisuppressor phenotype diagnostic for overexpressed release factors. The *SUP4* yeast strain used also carried the *ade2-1* ochre mutation, which causes adenine auxotrophy in a *SUP*$^+$ background. Suppression of the *ade2-1* mutation by *SUP4* tRNA ordinarily generates an adenine prototrophy (Ade$^+$) phenotype in a yeast strain of genotype *SUP4 ade2-1*; antisuppression of the *SUP4* tRNA by overexpression of any release factor (complex) should result in a reversion to adenine auxotrophy (Ade$^-$).

Overexpression of either the *SUP35* or *SUP45* genes in isolation had no antisuppressor effect on *SUP4*-mediated suppression, while overexpression of both *SUP35* and *SUP45* genes in the same strain produced a clear antisuppressor phenotype, indicated by the Ade$^-$ phenotype of the double transformant (Figure 3). This result indicates that Sup45p and Sup35p act in concert as components of a release factor complex to mediate translation termination in yeast, with neither protein alone being sufficient to outcompete the suppressor tRNA.

Xenopus and human SUP45 genes can couple effectively with yeast SUP35 to antisuppress an efficient tRNA suppressor

To confirm that Sup35p and Sup45p act together in a complex to mediate translation termination, a *SUP4 ade2-1* yeast strain carrying the yeast multicopy *SUP35* gene was transformed with a second plasmid expressing either the *Xenopus* or human *SUP45* (eRF1) homologues and tested for antisuppressor phenotype. The results show that overexpression of *Xenopus* eRF1 (Sup45p/Cl1) and *SUP35* generated an antisuppressor phenotype, as did overexpression of *SUP35* with human eRF1 (TB3-1; Figure 3). Neither overexpression of *Xenopus* eRF1 alone nor human eRF1 alone generated an antisuppressor phenotype (data not shown). The antisuppressor phenotype in the two heterologous combinations appeared qualitatively to be as efficient as that generated by the combined overexpression of yeast Sup45p and yeast Sup35p (Figure 3).

This result confirms that human and *Xenopus* eRF1 proteins can interact functionally with the yeast Sup35p protein producing a phenotype typical of that expected from release factor complex overexpression. The findings also lend further credence to the assertion that yeast Sup45p represents the *S.cerevisiae* eRF1 protein, since in multicopy the yeast *SUP45*, *Xenopus* Cl1 and human TB-3 genes all generate a similar antisuppression phenotype with the yeast *SUP35* gene.

Antisuppression due to co-expression of SUP35 and SUP45 is not restricted to a single suppressor tRNA species and is effective against ochre, amber and UGA suppression

To confirm that the simultaneous overexpression of *SUP35* and *SUP45* could act to antisuppress suppressor tRNAs

Table I. Simultaneous overexpression of Sup35p and Sup45p acts to antisuppress all three stop codons

Termination codon	Percentage stop codon readthrough[a]			Fold-reduction in suppressor efficiency[b]
	None (pEMBLYe23)	Sup45p (pVK62)	Sup45p and Sup35p (pVK63)	
UAA	4.6 ± 0.4	4.5 ± 0.34	1.77 ± 0.25	2.6
UAG	0.66 ± 0.02	0.65 ± 0.05	0.29 ± 0.04	2.3
UGA	0.23 ± 0.03	0.27 ± 0.06	0.18 ± 0.03	1.4

[a]Strain 5V-H19/DBY746 carrying the SUQ5 suppressor tRNA was transformed with either plasmid pEMBLYe23 (control), pVK62 to overexpress *SUP45* or pVK63 to overexpress both *SUP35* and *SUP45*, in different pairwise combinations with either pUKC815-L, pUKC817-L, pUKC818-L or pUKC819-L, to quantify nonsense suppression levels (Materials and methods). β-Galactosidase activities in each of the pUKC817-L, pUKC818-L or pUKC819-L transformants, representing the levels of UAA, UAG and UGA suppression respectively, were expressed as a percentage of the β-galactosidase activity in the pUKC815-L transformant. Values are the means of three independent assays ± 1 SD.
[b]Fold reduction in suppressor efficiency occuring with Sup45p and Sup35p co-overexpression.

other than *SUP4*, the experiment was repeated in a [*psi*$^+$] diploid strain (5V-H19/DBY746) heterozygous for *SUQ5*, which encodes the weak ochre suppressor tRNA^{Ser}. A [*psi*$^+$] strain was selected so as to elevate the ordinarily inefficient *SUQ5*-mediated suppression to more easily detectable levels (Cox, 1965).

To accurately quantify any antisuppressor effect, the plasmids pUKC815-L, pUKC817-L, pUKC818-L and pUKC819-L (Stansfield *et al.*, 1995; Materials and methods) were introduced into the *SUQ5*/*suq5*$^+$ [*psi*$^+$] strain overexpressing *SUP35* and *SUP45*. This was achieved by mating 5V-H19[pVK63] with DBY746 transformed with different pUKC-815 series vectors. pUKC815-L consists of a *PGK1–lacZ* gene fusion, while the pUKC817-L, pUKC818-L and pUKC819-L plasmids are identical to pUKC815-L except that one of the three termination codons, TAA, TAG and TGA respectively, is cloned in-frame at the junction of the *PGK1* and *lacZ* genes (Stansfield *et al.*, 1995). Any suppression of the in-frame premature stop codons will result in β-galactosidase activity and the levels of β-galactosidase activity can therefore be used to quantify suppressor or antisuppressor phenotypes.

In the [*psi*$^+$] *SUQ5*/*suq5*$^+$ strain the level of ochre codon suppression in a pUKC817-L transformant was 4.6% of control β-galactosidase activity of a pUKC815-L transformant (Table I). Overexpression of *SUP45* in a pVK62 transformant had no effect on stop codon suppression levels (Table I). However, simultaneous overexpression of *SUP35* and *SUP45* using plasmid pVK63 resulted in a 2.6-fold decrease in UAA suppression. The suppression of UAG and UGA termination codons observed in the [*psi*$^+$] *suq5*$^+$/*SUQ5* strain and measured using plasmids pUKC818-L and pUKC819-L was also reduced 2.3- and 1.4-fold respectively in the pVK63 transformants (Table I). We propose that this suppression derives from the action of natural suppressor tRNAs (Stansfield *et al.*, 1995), including a UGA suppressing tRNA^{Trp} (Tuite and McLauchlin, 1982), a UAG suppressing tRNA$^{Gln}_{CAG}$ (Pure *et al.*, 1988) and tRNA$^{Gln}_{CAA}$ capable of suppressing UAA (Weiss,W.A. and Frieberg, 1987).

To further confirm that the antisuppressor phenotype was effective against UGA suppression, the adenine prototroph strain MT576/5c, essentially *SUP4*^{UGA} *ade2*^{UGA}, was transformed with the different pair-wise combinations of plasmids detailed in Figure 4. These resulted variously in either *SUP35* or *SUP45* overexpression alone or simultaneous *SUP35* and *SUP45* overexpression. The results show that of all combinations only overexpression of Sup35p and Sup45p together is effective in antisuppressing the UGA suppressor tRNA, indicated by the adenine auxotrophy phenotype of this double transformant (Figure 4).

Taken together, the results presented in Table I and Figure 4 demonstrate clearly that overexpression of Sup35p and Sup45p together produces an antisuppressor phenotype effective against suppression at all three stop codons. This directly infers that the Sup35p–Sup45p protein complex acts to catalyse termination at all three stop codons.

The results additionally confirm that the antisuppressor phenotype produced when Sup35p and Sup45p are overexpressed together can act to antisuppress both a seryl-tRNA (*SUQ5*) and a tyrosyl-tRNA (*SUP4*^{UGA}) and that antisuppression is therefore not restricted to limiting the suppressor efficiency of a single tRNA species.

Discussion

Members of the *SUP45* (eRF1) gene family have release factor activity and can catalyse peptide chain release at all three stop codons (Frolova *et al.*, 1994), yet the eukaryote translation termination process is GTP dependent (Goldstein *et al.*, 1970; Konecki *et al.*, 1977). We have therefore proposed that *SUP35*, encoding a protein with C-terminal homology to translation elongation factor EF-1α, including concensus GTP binding sequence motifs (Kushnirov *et al.*, 1988; Wilson and Culbertson, 1988), might supply the necessary GTP hydrolytic activity (Tuite and Stansfield, 1994). Like *SUP45*, mutant alleles of the *SUP35* gene can be isolated with either an omnipotent suppressor or allosuppressor phenotype (Inge-Vechtomov and Andrianova, 1970; Cox, 1977), an indication that Sup35p may play a role in translation termination. It is also likely that homologues of the *SUP35* gene will be found in a wide variety of organisms; for example a human cDNA has been isolated encoding a protein with 52.3% identity to *S.cerevisiae* Sup35p (Hoshino *et al.*, 1989) and *SUP35* homologues have been identified in the yeast *Pichia pinus* (Kushnirov *et al.*, 1990) and in *Xenopus laevis* (Zhouravleva *et al.*, 1995)

We show here, using the two-hybrid system, that the yeast eRF1(Sup45p) protein interacts *in vivo* with Sup35p (Figure 1) This interaction was also demonstrated *in vitro*

Double Transformant			Relevant protein(s) expressed		Three independent double-transformants grown on defined medium (-ura -leu)
					+adenine / -adenine
pRS425	+	YEp24	-	-	
pUKC606	+	YEp24	Sup35p(Sc)	-	
pRS425	+	pUKC802	-	Sup45p(Sc)	
pUKC606	+	pUKC802	Sup35p(Sc)	+ Sup45p(Sc)	

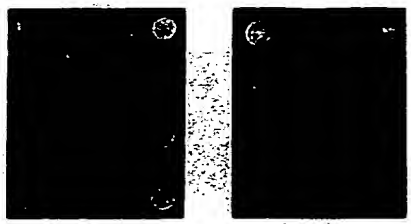


Fig. 4. Simultaneous overexpression of Sup35p and Sup45p also acts to antisuppress the UGA suppressor tRNA $SUP4^{UGA}$. Strain MT576/5c carrying the $SUP4^{UGA}$ tRNA suppressor and the $ade2^{UGA}$ allele was transformed with the pair-wise combinations of plasmids indicated and growth of three independent transformants on the control medium (+ adenine) compared with growth on defined medium lacking adenine. Each yeast 'colony' represents ~1×10^5 cells spotted onto the defined medium in 3 μl water and allowed to grow for 3 (+ adenine plate) or 5 days (-adenine plate) at 30°C.

using immobilized Sup45p to precipitate Sup35p (Figure 2). We have subsequently used gel filtration to demonstrate that Sup35p and Sup45p exist as a heterodimer in yeast cell lysates (data not shown), again confirming that the two proteins form a complex *in vivo*. That this interaction is functionally significant is demonstrated by the anti-suppression phenotype accompanying simultaneous over-expression of the *SUP35* and *SUP45* genes (Figures 3 and 4): this phenotype is typical of and diagnostic for release factor genes, having been demonstrated for the *prfA* and *prfB* genes encoding *E.coli* release factors RF1 and RF2 respectively (Weiss,R.B. *et al.*, 1984) and the *S.cerevisiae* *MRF1* gene encoding the mitochondrial release factor mRF1 (Pel *et al.*, 1992). In each case increasing the cellular levels of the release factor in relation to the suppressor tRNA acts to out-compete the tRNA species for stop codon binding. Furthermore, we show here that this antisuppression phenotype is effective against UAA, UAG and UGA suppressors (Table I and Figure 4), inferring that Sup35p and Sup45p together form an omnipotent release factor complex. This finding corroborates the results of genetic studies, where mutations in either the *SUP35* or *SUP45* gene resulted in an omnipotent suppressor phenotype (Inge-Vechtomov and Andrianova, 1970; Hawthorne and Leupold, 1974; Stansfield *et al.*, 1995). The omnipotence of the yeast release factor complex, observed *in vivo*, parallels the *in vitro* situation in *Xenopus*, where Sup35p is able to enhance, in a GTP-dependent manner, the peptidyl release activity of Sup45p at all three stop codons (Zhouravleva *et al.*, 1995).

The case for Sup35p and Sup45p together forming a release factor complex is strengthened by our finding that either of the higher eukaryote *SUP45* homologues, *Xenopus* C11 and human TB3-1, both with demonstrated eRF1 activity (Frolova *et al.*, 1994), can act in this putative release factor complex and couple functionally with yeast Sup35p, as indicated by their ability to act as multicopy antisuppressors in a strain overexpressing *SUP35* (Figure 3).

While *Xenopus* and human eRF1 (Sup45p) proteins have demonstrated release factor activity *in vitro* (Frolova *et al.*, 1994), it is not clear whether the level of termination activity demonstrated in the *in vitro* assay would be sufficient to out-compete any nonsense suppressor tRNA cognate for one of the stop codons if added to the assay, i.e. can Sup45p (eRF1) mediate efficient antisuppression alone *in vitro*? We would argue not, based on the results presented here, proposing instead that Sup45p is necessary, but not sufficient, for efficient *in vivo* termination. Rather, both the Sup45p and Sup35p proteins are required to generate the levels of release factor complex activity necessary to out-compete suppressor tRNAs and for normal cellular translation.

Confirmation that Sup35p represents a second component of the eukaryote release factor comes from evidence that in the *in vitro* termination assay Sup35p can stimulate eRF1 (Sup45p) activity in a GTP-dependent manner (Zhouravleva *et al.*, 1995). Nevertheless, the findings presented here raise questions about the mechanism of eukaryotic translation termination and its relationship to the corresponding prokaryote process, as defined by studies with *E.coli*.

A number of differences between the two are immediately apparent. Sup35p could, with GTP binding capability, perform a role analogous to that of the prokaryotic termination factor RF3. RF3, encoded by a non-essential gene, stimulates RF1- and RF2-mediated activity in a GTP-dependent manner (Milman *et al.*, 1969) and shows sequence homology to elongation factor EF-G (equivalent to the eukaryote elongation factor EF-2; Grentzmann *et al.*, 1994; Mikuni *et al.*, 1994). In contrast, Sup35p is encoded by an essential gene and shows considerable C-terminal homology to elongation factor EF-1α (Kushnirov *et al.*, 1988; Wilson and Culbertson, 1988). The homology of RF3 and Sup35p to different elongation factors argues for a fundamental mechanistic difference between the prokaryote and eukaryote termination processes. Perhaps as important in this respect is the observation that over-expression of both Sup45p and Sup35p is required to enhance the efficiency of termination (Figure 3), again a point of difference with prokaryotes, where overexpression of RF1 or RF2 alone is sufficient to generate an antisuppressor phenotype (Weiss,R.B. *et al.*, 1984).

The identification of Sup35p as an essential protein interacting with Sup45p (eRF1) is thus an important first step in the characterization of the components of the

Table II. Plasmids used in this study

Plasmid[a]	Protein expressed[b]	Source
pUKC802	Sup45p	Akhmaloka (1991)
pUKC606	Sup35p	This study
pVK62	Sup45p	This study
pVK63	Sup45p and Sup35p	This study
pEX1	*Xenopus* eRF1 (Sup45p) (yeast Sup45p1–10, *Xenopus*, 11–437)	M.Philippe
pEO6	Human eRF1 (Sup45p) (yeast Sup45p1–10, human, 11–437)	M.Kress
Two-hybrid vectors		
pUKC601	Gal4p bd (1–147)–Sup45p(10–437)	This study
pUKC605	Gal4p ad (768–881)–Sup35p(41–685)	This study
pG–DD	Gal4p ad (768–881)–Cpf1p(266–351)	J.Mellor
p13–DD	Gal4p bd (1–147)–Cpf1p(266–351)	J.Mellor
pGBT9	Gal4p bd (1–147)	S.Fields
pGAD424	Gal4p ad (768–881)	S.Fields

[a]All plasmids in the table carry the 2μ origin of replication and are therefore multicopy in yeast.
[b]Figures in brackets refer to the amino acids of each protein expressed, numbered with the initiator methionine residue as 1.

eukaryote termination process and will expedite further definition of the mechanisms involved. The findings presented here also have a direct bearing on related and developing issues, such as the suggestion that Sup35p may be a prion-type protein, able to adopt different conformations: it is proposed that each conformation exhibits distinct activities with respect to translation termination, giving rise to the yeast non-Mendelian genetic phenomenon [*psi*] (Cox, 1994; Tuite, 1994; Wickner, 1994). Finally, we forsee further interest in the notion of regulation of translation at the level of termination, with the recent finding that the *SAL6* gene, mutant alleles of which can act to enhance the activity of *sup35* and *sup45* omnipotent suppressors (Song and Liebman, 1987), codes for a protein, Sal6p, with homology to serine/threonine phosphatases (Vincent *et al.*, 1994).

Materials and methods

Strains and media

The *S.cerevisiae* strains used in this study were: MT552/36d (MATα *ura3-1*); MT422/1c (MATα *ura3-1 leu2-3,112 his5-2 ade2-1 SUP4-o*); 5V-H19 (MATa *ade2-1 can1-100 leu2-3,112 ura3-52 SUQ5* [*psi*$^+$]); DBY746 (MATα *his3-1 leu2-3,112 ura3-52 trp1-289*); MT576/5c (MATα *SUP4*^{UGA} *ade2*^{UGA} *ura3-1 leu2-3,112*); Y526 (MATa *ura3-52 his3-200 ade2-101 lys2-801 trp1-901 leu2-3,112 can-1 gal4-542 gal80-538 ura3::GAL1-lacZ*). Y526 (Bartel *et al.*, 1993a) was a gift from S.Fields (State University of New York at Stony Brook). *Escherichia coli* strain DH5α (F'*lendA1 hsdR17* ($r_K^-m_K^+$) *supE44 thi-1 recA1 gyrA* (Nal^R) *relA1* Δ*lacZYA-argF*) U169 (ϕ80*dlac* Δ(*lacZ*)*M15*) was used for all cloning experiments. Yeast cultures were grown using standard conditions (Sherman, 1991) in YEPD liquid medium (2% w/v Bactopeptone, 1% w/v yeast extract and 3% w/v glucose). Yeast transformants were grown in 2% w/v glucose, 0.67% w/v yeast nitrogen base without amino acids (Difco), supplemented with the required amino acids and co-factors. Bacteria were grown in LB broth (Sambrook *et al.*, 1989).

Plasmid construction and DNA manipulation

All DNA manipulations and plasmid construction techniques were carried out using standard protocols (Sambrook *et al.*, 1989). The salient features of the plasmids used in this work are outlined in Table II.

Plasmid pEX1 consists of the yeast *SUP45* promoter and the first 10 codons of the yeast *SUP45* coding sequence ligated to the C-terminal 428 codons of the *Xenopus SUP45* (eRF1) gene. Plasmid pEO6 is identical, except the sequence coding for the C-terminal 427 codons of the human eRF1 (*SUP45*) gene replaces the *Xenopus* sequence of pEX1. These plasmids were generously donated by Dr M.Kress (Laboratoire d'Oncologie Moléculaire, Villejuif, France) and Dr M.Philippe (University of Rennes, Rennes, France). Plasmid pUKC606 consists of the yeast *SUP35* coding sequence and promoter on an *Xho*I–*Not*I fragment derived from plasmid pSM138 (Doel *et al.*, 1994) ligated into the multicopy vector pRS425 (Christiansen *et al.*, 1992) cut with *Xho*I and *Not*I. Plasmid pUKC802 was made by cloning the yeast *SUP45* promoter and gene into the multicopy vector YEp24 (Akhmaloka, 1991). Plasmid pUKC601 is a derivative of pGBT9 (Bartel *et al.*, 1993b) and consists of the yeast *SUP45* gene sequence coding for amino acids 10–437 cloned in-frame and downstream of the *GAL4* binding domain. pUKC601 was constructed by cloning the *SUP45* coding sequence and promoter on a *Sal*I–*Xho*I fragment cut from pUKC802 (Akhmaloka, 1991) into pBluescript IIKS+ cut with *Sal*I, generating pUKC600. The *SUP45* sequence coding for amino acids 10–437 was excised from pUKC600 on a *Bgl*II–*Sal*I fragment and ligated into pGBT9 cut with *Bam*HI and *Sal*I. Plasmid pUKC605 is derived from pGAD424 (Bartel *et al.*, 1993b) and was constructed by cloning the *SUP35* coding sequence and promoter on a *Xba*I fragment excised from pSM138 (Doel *et al.*, 1994) into plasmid pSP73 (Promega Corporation) cut with *Xba*I, generating pSP73SUP35. The *SUP35* coding sequence representing amino acids 41–685 was excised from pSP73SUP35 on a *Pst*I fragment and ligated into pGAD424 cut with *Pst*I, downstream of and in-frame with the *GAL4* activation domain, creating pUKC605.

Plasmid pUKC625, a derivative of pUKC802 (Akhmaloka, 1991), contains the *SUP45* gene with an adaptor inserted between the *Bam*HI and *Xho*I sites coding for a C-terminal affinity tag of six histidine residues.

```
5'-GATCCAGACATCACCATCACCATCACTAAGCTTGAGC-3'
    3'-GTCTGTAGTGGTAGTGGTAGTGATTCGAACTCGAGCT-5'
```

Plasmids p13-DD and pG-DD consist of the DNA sequence coding for the dimerization domain (amino acids 266–351) of transcription factor *CPF1* cloned in an in-frame fusion with either the *GAL4* binding or activation domains (Dowell *et al.*, 1992); both p13-DD and pG-DD were generous gifts of Dr Jane Mellor (University of Oxford, UK).

Plasmid pVK62 was constructed by ligating the *Xba*I–*Bam*HI fragment from pYsup1-1 (Breining *et al.*, 1984) into pEMBLYE23 (Cesareni and Murray, 1987) cleaved using *Xba*I and *Bcl*I. Plasmid pVK63 was constructed by cloning the *Bgl*II–*Bam*HI fragment from pSTR4 (Telckov *et al.*, 1986) containing the complete *SUP35* coding sequence and promoter into pVK62 cut with *Bam*HI.

Plasmids pUKC815-L, pUKC817-L, pUKC818-L and pUKC819-L were used in this work and are derived from pUKC815, pUKC817, pUKC818 and pUKC819 respectively (Stansfield *et al.*, 1995). The 'L' in the pUKC815-L series vectors used in this work designates replacement of the plasmid selectable marker *URA3* on a *Sma*I–*Sal*I fragment with the *LEU2* gene on a *Sma*I–*Sal*I fragment derived from plasmid pJJ250 (Jones and Prakash, 1990).

β-Galactosidase assays of yeast strains transformed with the pUKC815/817/818/819 series vectors

β-Galactosidase assays were performed according to the method of Finkelstein and Strausberg (1983).

Yeast transformation with plasmid DNA

Yeast were transformed by electroporation according to the method of Becker and Guarente (1991), using BioRad Gene-pulser equipment according to the instructions of the manufacturer.

Preparation of post-ribosomal supernatants

Yeast cultures were grown to a cell density of 3.5×10^7 cells/ml. Harvested cells were washed with lysis buffer (25 mM Tris–HCl, pH 7.2, 5 mM $MgCl_2$, 5 mM 2-mercaptoethanol and 25 mM KCl) containing 1 mM phenylmethylsulphonyl fluoride (PMSF), 1 mM benzamidine, 10 μM leupeptin and 10 μM pepstatin A to limit proteolytic degradation. The concentration of KCl in the lysis buffer was increased to 0.8 M if the PRS was to be subsequently used in the purification of Sup45p$(His)_6$. Following washing in lysis buffer, harvested cells were resuspended in a minimum volume of the same buffer. Cells were lysed by vortexing with glass beads and cell debris removed by centrifugation at 13 000 r.p.m. for 15 min in a benchtop microcentrifuge, producing a post-mitochondrial supernatant (PMS). The PMS was centrifuged in a Beckman TL100.3 rotor at 50 000 r.p.m. for 75 min, resulting in the sedimentation of polysomal ribosomes, 80S ribosomes and 40S and 60S subunits, with the generation of a PRS.

Immobilization of Sup45p$(His)_6$

The following operations were performed at room temperature unless otherwise stated and columns were allowed to flow by gravity.

A 500μl column of Ni–NTA resin (Qiagen) was equilibrated with 30 mM bis-Tris–HCl, pH 6.8, 0.8 M KCl, 5 mM 2-mercaptoethanol, 30 mM imidazole (buffer A). The PRS containing Sup45p$(His)_6$ was passed three times through the resin, which was then washed with 10 column vol. buffer A followed by 10 vol. buffer containing 30 mM bis-Tris–HCl, pH 6.2, 1 M KCl, 5 mM 2-mercaptoethanol, 20% glycerol, 30 mM imidazole. This procedure resulted in partial purification and immobilization of Sup45p$(His)_6$ on the Ni–NTA resin.

In order to assess the binding of Sup35p to the immobilized Sup45p$(His)_6$, the resin was equilibrated with 10 column vol. lysis buffer containing 25 mM KCl. A PRS prepared from cells overexpressing Sup35p (transformed with plasmid pUKC606) was then incubated with a suspension of the resin in 25 mM KCl lysis buffer for 2 h on a shaking platform at 4°C. After this time the resin was centrifuged briefly, the supernatant removed and retained and the resins returned to columns for washing with 10 column vol. 25 mM KCl lysis buffer containing 0.15% w/v Tween-20 detergent.

Protein gel electrophoresis and Western blot analysis

SDS–polyacrylamide gel electrophoresis (SDS–PAGE) was performed according to standard protocols (Laemmli, 1970). SDS–PAGE gels contained 10% w/v acrylamide. Western blotting onto nitrocellulose was performed according to standard protocols (Towbin *et al.*, 1979; Harlow and Lane, 1988). Western blots were probed with either an anti-Sup45p polyclonal rabbit antibody used at 1:100 dilution (Stansfield *et al.*, 1992) or a polyclonal rabbit anti-Sup35p antibody used at 1:2000 dilution (Didichenko *et al.*, 1991). Bound antibody was detected using the Amersham ECL system according to the manufacturer's instructions.

Identification and quantification of two-hybrid system protein–protein interactions

Strain Y526 (Bartel *et al.*, 1993a), transformed with pair-wise combinations of two-hybrid vectors, was tested for β-galactosidase activity initially using X-gal staining of yeast colonies grown on nitrocellulose laid on agar medium, according to the method of Breeden and Nasmyth (1985). β-Galactosidase activities were further quantified by growing the transformants in liquid medium to a cell density of 4×10^7 cells/ml. A cell lysate was prepared in Z-buffer (Finkelstein and Strausberg, 1983) containing 1 mM phenylmethylsulphonyl fluoride, using glass bead lysis. Lysates were made up to 1 ml using Z-buffer and 200 μl *o*-nitrophenyl galactoside (ONPG; 4 mg/ml in water) added. The enzyme reaction rate was monitored at 420 nm and enzyme activity calculated using an extinction coefficient for ONP of 0.0045 nmol/ml. β-Galactosidase activities were expressed as pmol ONP produced/min/mg protein. Protein concentrations were determined using standard protocols (Lowry *et al.*, 1951).

Acknowledgements

This work was supported by a Biotechnology and Biological Sciences Research Council (BBSRC) grant to M.F.T., a BBSRC grant to B.S.C., a Wellcome Trust grant to V.V.K. and by grants from the Fogarty Foundation and the Russian Foundation for Basic Research to M.D.T.A. We thank Stan Fields and Paul Bartel (State University of New York at Stony Brook) for the generous gift of plasmids pGBT9 and pGAD424 and the yeast strain Y526, Michel Philippe (University of Rennes) and Michel Kress (Laboratoire d'Oncologie Moléculaire, Villejuif, France) for the generous gift of plasmids pEX1 and pEO6 respectively and Jane Mellor and Jimmy Tsang (University of Oxford) for the gift of plasmids pDD-13 and pG-DD.

References

Akhmaloka (1991) PhD Thesis, University of Kent, Canterbury, UK.
Bartel,P.L., Chien,C.T., Sternglanz,R. and Fields,S. (1993a) *Biotechniques*, **14**, 920–924.
Bartel,P.L., Chien,C.T., Sternglanz,R. and Fields,S. (1993b) In Hartley,D.A. (ed.), *Cellular Interactions in Development: A Practical Approach*. Oxford University Press, Oxford, UK, pp. 153–179.
Becker,D.M. and Guarente,L. (1991) *Methods Enzymol.*, **194**, 182–186.
Breeden,L. and Nasmyth,K. (1985) *Cold Spring Harbor Symp. Quant. Biol.*, **50**, 643–650.
Breining,P., Surguchov,A.P. and Piepersberg,W. (1984) *Curr. Genet.*, **8**, 467–470.
Caskey,C.T., Tompkins,R., Scolnick,E., Caryk,T. and Niremberg,M. (1968) *Science*, **162**, 135–138.
Caskey,C.T., Beaudet,A.L. and Tate,W.P. (1973) *Methods Enzymol.*, **30F**, 293–303.
Cesareni,G. and Murray,A.H. (1987) In Setlow,J.K. (ed.), *Genetic Engineering: Principles and Methods*. Plenum Press, New York, NY, Vol. 4, pp. 135–154.
Chernoff,Y.O., Derkach,I.L. and Inge-Vechtomov,S.G. (1993) *Curr. Genet.*, **24**, 268–270.
Chien,C.-T., Bartel,P.L., Sternglanz,R. and Fields,S. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 9578–9582.
Christianson,T.W., Sikorski,R.S., Dante,M., Shero,J.H. and Heiter,P. (1992) *Gene*, **110**, 119–122.
Cox,B.S. (1965) *Heredity*, **20**, 505–521.
Cox,B.S. (1977) *Genet. Res.*, **30**, 187–205.
Cox,B.S. (1994) *Curr. Biol.*, **4**, 744–748.
Didichenko,S.A., Ter-Avanesyan,M.D. and Smirnov,V.N. (1991) *Eur. J. Biochem.*, **198**, 705–711.
Doel,S.M., McCready,S.J., Nierras,C.R. and Cox,B.S. (1994) *Genetics*, **137**, 659–670.
Dowell,S.J., Tsang,J.S.H. and Mellor,J. (1992) *Nucleic Acids Res.*, **20**, 4229–4236.
Finkelstein,D.B. and Strausberg,S. (1983) *Mol. Cell. Biol.*, **3**, 1625–1633.
Fields,S. and Song,O. (1989) *Nature*, **340**, 245–246.
Frolova,L.Y., Dalphin,M.E., Justesen,J., Powell,R.J., Drugeon,G., McCaughan,K.K., Kisselev,L.L., Tate,W.P. and Haenni,A.-L. (1993) *EMBO J.*, **12**, 4013–4019.
Frolova,L. *et al.* (1994) *Nature*, **372**, 701–703.
Goldstein,J.L., Beaudet,A.L. and Caskey,C.T. (1970) *Proc. Natl Acad. Sci. USA*, **67**, 99–106.
Grenett,H.E., Bounelis,P. and Fuller,G.M. (1992) *Gene*, **110**, 239–243.
Grentzmann,G., Brechemier-Baey,D., Heurgue,V., Mora,L. and Buckingham,R.H. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 5848–5852.
Harlow,E. and Lane,D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Hawthorne,D.C. and Leupold,U. (1974) *Curr. Topics Microbiol. Immunol.*, **64**, 1–47.
Himmelfarb,H.J., Maicas,E. and Friesen,J.D. (1985) *Mol. Cell. Biol.*, **5**, 816–822.
Hirashima,A. and Kaji,A. (1972) *J. Mol. Biol.*, **65**, 43–58.
Hoshino,S., Miyazawa,H., Enomoto,T., Hanaoka,F., Kikuchi,Y. and Kikuchi,A.U.M. (1989) *EMBO J.*, **8**, 3807–3814.
Inge-Vechtomov,S.G. and Andrianova,V.M. (1970) *Genetika*, **6**, 103–115.
Jones,J.S. and Prakash,L. (1990) *Yeast*, **6**, 363–366.
Klein,H.A. and Capecchi,M.R. (1971) *J. Biol. Chem.*, **246**, 1055–1061.
Konecki,D.S., Aune,K.C., Tate,W. and Caskey,C.T. (1977) *J. Biol. Chem.*, **252**, 4514–4520.
Kushnirov,V.V., Ter-Avanesyan,M.D., Telckov,M.V., Surguchov,A.P., Smirnov,V.N. and Inge-Vechtomov,S.G. (1988) *Gene*, **66**, 45–54.
Kushnirov,V.V. *et al.* (1990) *Yeast*, **6**, 461–477.
Laemmli,U.K. (1970) *Nature*, **227**, 680–689.
Lee,C.C., Craigen,W.J., Muzny,D.M., Harlow,E. and Caskey,C.T. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 3508–3512.
Lowry,O.H., Rosebrough,N.J., Farr,A.L. and Randall,R.J. (1951) *J. Biol. Chem.*, **193**, 265–275.
Milman,G., Goldstein,J., Scolnick,E. and Caskey,C.T. (1969) *Proc. Natl Acad. Sci. USA*, **63**, 183–190.
Mikuni,O., Ito,K., Moffat,J., Matsumara,K., McCaughan,K.,

Nobukuni,T., Tate,W. and Nakamura,Y.(1994) *Proc. Natl Acad. Sci. USA*, **91**, 5798–5802.
Pel,H.J., Maat,C., Rep,M. and Grivell,L.A. (1992) *Nucleic Acids. Res.*, **20**, 6339–6346.
Pure,G.A., Robinson,G.W., Naumovski,L. and Frieberg,E.C. (1988) *J. Mol. Biol.*, **183**, 31–42.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Scolnick,E., Tompkins,R., Caskey,C.T. and Niremberg,M. (1968) *Proc. Natl Acad. Sci. USA*, **61**, 768–774.
Sherman,F. (1991) *Methods Enzymol.*, **194**, 3–20.
Song,J.M. and Liebman,S.W. (1987) *Genetics*, **115**, 451–460.
Stansfield, I., Grant, C.M., Akhmaloka and Tuite, M.F. (1992) *Mol. Microbiol.*, **6**, 3469–3478.
Stansfield,I., Akhmaloka and Tuite,M.F. (1995) *Curr. Genet.*, **27**, 417–426.
Tassan,J.P., Le Guellec,K., Kress,M., Faure,M., Camonis,J., Jacquet,M. and Philippe,M. (1993) *Mol. Cell. Biol.*, **13**, 2815–2821.
Telckov,M.V., Surguchov,A.P., Dagkesamanskaya,A.R. and Ter-Avanesyan,M.D. (1986) *Genetika*, **22**, 17–25.
Ter-Avanesyan,M.D., Inge-Vechtomov,S.G. and Surguchov,A.P. (1984) *Dokl. Akad. Nauk SSSR*, **277**, 713–716.
Ter-Avanesyan,M.D., Didichenko,S.A., Kushnirov,V.V. and Dagkesamanskaya,A.R. (1993) In Brown,A.J.P., Tuite,M.F. and McCarthy,J.E.G. (eds), *Protein Synthesis and Targeting in Yeast*. Springer-Verlag, Heidelberg, Germany, NATO ASI series Vol. 71, pp. 81–90.
Towbin,H., Staehelin,T. and Gordon,J. (1979) *Proc. Natl Acad. Sci. USA*, **76**, 4350–4354.
Tuite,M.F. (1994) *Nature*, **370**, 327–328.
Tuite,M.F. and McLaughlin,C.S. (1982) *Mol. Cell. Biol.*, **2**, 490–497.
Tuite,M.F. and Stansfield,I. (1994) *Nature*, **372**, 614–615.
Vincent,A., Newnam,G. and Liebman,S.W. (1994) *Genetics*, **138**, 597–607.
Weiss,R.B., Murphy,J.P. and Gallant,J.A. (1984) *J. Bacteriol.*, **158**, 362–364.
Weiss,W.A. and Frieberg,E.C. (1987) *J. Mol. Biol.*, **192**, 725–735.
Wickner,R.B. (1994) *Science*, **264**, 566–569.
Wilson,P.G. and Culbertson,M.R. (1988) *J. Mol. Biol.*, **199**, 559–573.
Zhouravleva,G., Frolova,L., Le Goff,X., Le Guellec,R., Inge-Vechtomov,S., Kisselev,L. and Philippe,M. (1995) *EMBO J.*, **14**, 4065–4072.

Received on February 3, 1995; revised on June 16, 1995

Volume 215, number 2, 257-260 FEB 04683 May 1987

Localization of possible functional domains in *sup2* gene product of the yeast *Saccharomyces cerevisiae*

Vitaliy V. Kushnirov, Michail D. Ter-Avanesyan, Andrei P. Surguchov, Vladimir N. Smirnov and Sergey G. Inge-Vechtomov*

Institute of Experimental Cardiology, USSR Cardiology Research Center, 3rd Cherepkovskaya Str., 15a, 121552 Moscow and **Department of Genetics, Leningrad State University, Universitetskaya nab., 7/9, 199164 Leningrad, USSR*

Received 25 February 1987

Primary structures of yeast *sup2* gene and polypeptide product coded by the gene are compared with the current nucleotide and amino acid sequence data base. The amino acid sequence of the *sup2* product shows homology to elongation factors from different sources. Especially high homology is found in the regions, corresponding to conservative aminoacyl-tRNA- and GTP-binding domains, described in elongation factors and other proteins. The data obtained are discussed in relation to the functions of *sup2* polypeptide product in protein synthesis.

Protein synthesis; GTP-binding site; aminoacyl-tRNA-binding site; Gene homology

1. INTRODUCTION

In the yeast *Saccharomyces cerevisiae* in addition to well known dominant suppressors coding for tRNA a new class of recessive suppressors has been described [1–4]. These suppressors designated *sup1* and *sup2* (similar to *sup45* and *sup35*, respectively [4,5]) were found to be omnipotent, acting towards all three types of nonsense mutations (UAG, UGA and UAA), the suppression being mediated by an increase in the translational ambiguity [4,6,7]. These data indicate that *sup1* and *sup2* genes code for proteins controlling the accuracy of codon-anticodon interaction. Although the functional properties of these suppressors are well characterized [4,6,7] and both genes have recently been cloned [5,8–10], the opinion about the nature of their polypeptide products remains controversial. It seems that they combine the properties of ribosomal proteins [4,6,7] and protein factors [5,8] affecting different parameters of translation in yeast, in particular, the level of fidelity.

Here we present the results of a computer-assisted comparison of the *sup2* gene polypeptide product with published sequences, allowing us to find considerable homology to elongation factors (EFs). Homologous regions include several domains in EFs, for which a functional role has been proposed earlier.

Correspondence address: A.P. Surguchov, Institute of Experimental Cardiology, 3rd Cherepkovskaya Str., 15a, 121552 Moscow, USSR

2. MATERIALS AND METHODS

The cloning strategy for *sup2* gene is described earlier [10]. Nucleotide sequence was determined following Sanger et al. [11]. The complete sequence of the *sup2* gene and flanking regions will be published elsewhere. Primary structures were compared using the program GENEUS [12].

3. RESULTS

The search of nucleotide sequences homologous to the *sup2* gene in the EMBL data bank (5789 se-

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

quences) and further analysis on the amino acid level indicated the existence of homology of the *sup*2 gene product with yeast EF-1α and with analogous EFs from other species, mitochondria and chloroplasts. The highest level of homology to *sup*2 gene product was found for EF-1α from yeast [13,14] and brine shrimp, *Artemia salina* [15]. This homology spans the full length of EFs and permits an alignment of the *sup*2 gene product to either protein from EF-1α family.

A region of the *sup*2 gene product, homologous to EFs corresponds to a part of the open reading frame of the gene, starting from the third in-frame ATG codon to the termination codon. However, there are indications that this region may represent a functionally active protein. For example, plasmids, carrying the *sup*2 gene, in which initiation of translation on the first and second in-frame ATG codons is impaired due to deletion, retain the ability to complement a temperature-sensitive

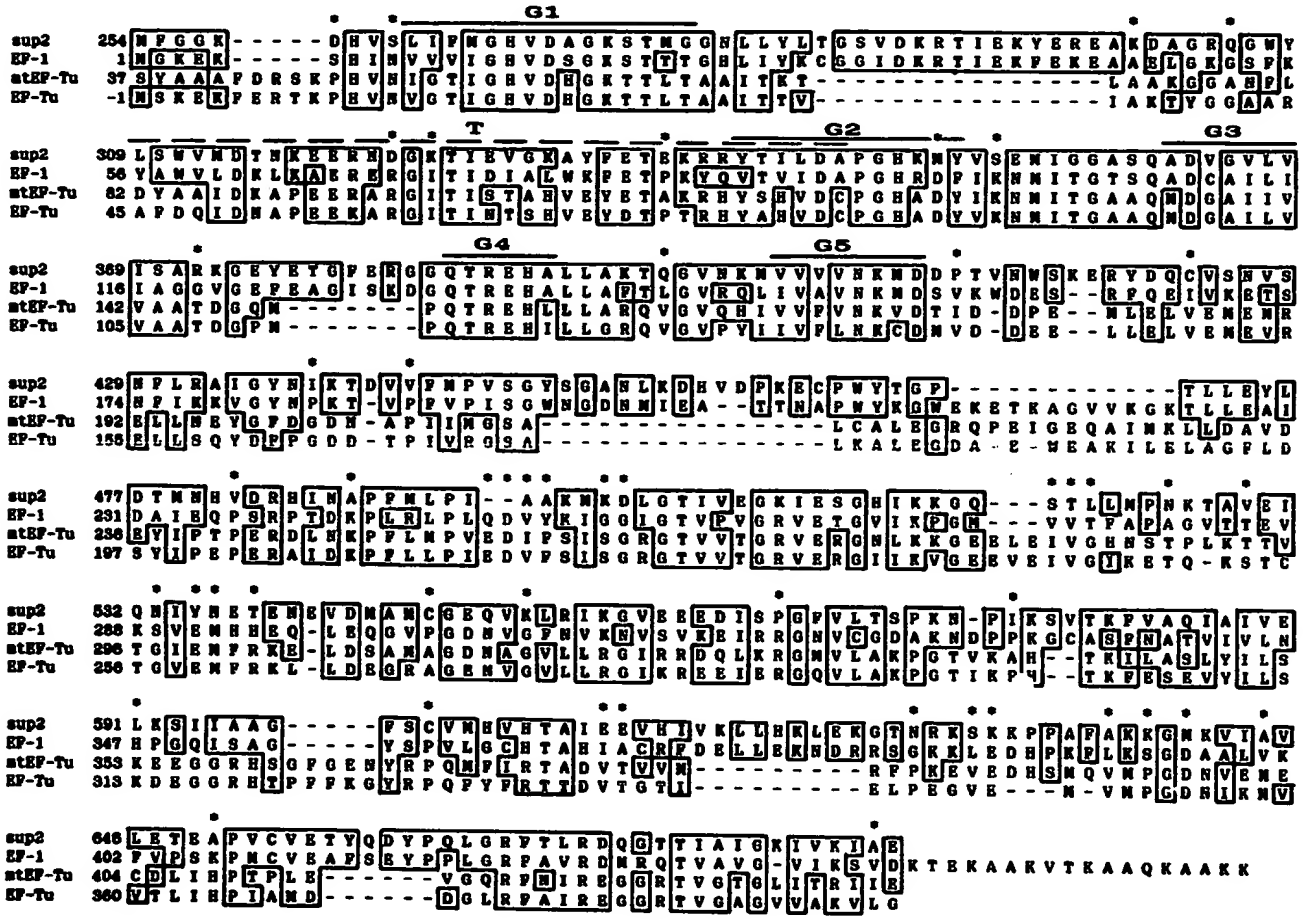


Fig.1. Comparison of the amino acid sequence of *sup*2 polypeptide product, yeast EF-1, mtEF-Tu and *E. coli* EF-TuA. The amino acid sequences of yeast *sup*2 polypeptide product, EF-1α [13], mtEF-Tu [16] and *E. coli* EF-TuA [17] are aligned to give maximal homology by introducing several gaps (−). The one-letter amino acid notation is used. The amino acid residue number 1 in the *sup*2 product is tentatively assigned to methionine at the first codon ATG in the open reading frame while that for EF-1α and mtEF-Tu to methionine at the initiator codon ATG and that for *E. coli* EF-TuA to serine, which is located at the N-terminal of the protein. The regions of exact homology and conservative substitutions between the *sup*2 product and either elongation factor are indicated by boxes. Conservative domains G_1–G_5 (involved in GTP-binding [19]) are indicated by solid lines, whereas region T (important for aminoacyl-tRNA binding) is shown by a dashed line. Positions, where EFs are homologous between themselves, but non-homologous to the *sup*2 product are marked by an asterisk (*). The following Dayhoff conservative categories [18] were used: C; S, T, P, A, G; N, D, E, Q; H, R, K; M, I, L, V; and F, Y, W.

mutation in the *sup*2 gene (Telckov, M., personal communication).

In fig.1 the *sup*2 gene product amino acid sequence, starting from the third methionine to the C-terminus (amino acids 254–685), is aligned to the sequences of yeast EF-1α [13], mitochondrial EF-Tu [16] and *E. coli* EF-TuA [17]. Comparison of the three latter sequences, belonging to evolutionary distant sources, reveals the most conserved regions of the EF-1α family. As seen from fig.1, most of them are present in the *sup*2 gene product sequence, although in some cases, amino acids conserved in the three proteins correspond to non-homologous amino acids in the *sup*2 gene product (shown by asterisks). Considering conservative amino acid substitutions [18] as homologous and without counting the gaps, the sequence of the *sup*2 gene product shows 62% homology with yeast EF-1α and 43% with yeast mitochondrial EF-Tu. For comparison, homology between yeast cytosolic EF-1α and mitochondrial EF-Tu amounts to 55% [13].

From fig.1 one can see that the degree of homology is distributed unevenly along four sequences, the highest homology being in the N-terminus of EFs, which is where the conservative domains are located for which a functional role has been proposed [19–22]. Earlier the comparison of primary structures of several GTP-binding proteins, e.g. EFs, bacterial initiation factor IF-2α, *ras* proteins and bovine transducin, allowed deduction the structure of the GTP-binding site [19], including five conservative domains located sequentially. As shown in fig.1, a similar structural organisation is present in the *sup*2 gene product.

A functional role for another conservative domain in EFs was elucidated in the experiments on chemical modification and photooxidation of *E. coli* EF-Tu. A stretch of amino acids between 44 and 81 was shown to be involved in aminoacyl-tRNA binding [20–22]. A corresponding region homologous to this aminoacyl-tRNA binding domain is located between amino acids 308 and 345 in the *sup*2 product sequence (fig.1).

4. DISCUSSION

Upon alignment of the *sup*2 gene product with yeast EF-1α one can see rather high homology throughout almost the entire sequence. The predicted secondary structure of these proteins (α-helix, β-sheets, β-turns) as well as hydrophilicity distribution are very similar (not shown). This could mean that these two proteins may have similar tertiary structure and interact with common ligands. A significant homology found between a polypeptide product of the *sup*2 gene and aminoacyl-tRNA and GTP-binding domains in EFs of different origin may indicate that the amino acid sequences shown in fig.1 (G_1, G_2, G_3, G_4, G_5 and T) are specialized for performing the same functions in the *sup*2 gene product (GTP-binding, GTP-hydrolysis and aminoacyl-tRNA recognition).

These data together with a previous observation on the participation of the *sup*2 gene in the control of translational fidelity [4,6,7] allow us to suggest that a polypeptide product of the gene may perform GTP-dependent proofreading of codon-anticodon interaction in the ribosome acceptor site. The presence of structures homologous to aminoacyl-tRNA binding domain indicates that the *sup*2 gene product may directly participate in the process of aminoacyl-tRNA recognition.

It is important to note that participation in the control of translational fidelity is already proven for one of the proteins homologous to the *sup*2 gene product. In vitro studies of mutationally altered *E. coli* EF-Tu, namely EF-Tu Ar reveal that it increases the errors at both the proofreading and the initial aminoacyl-tRNA selection steps [23]. This mutation together with mutation inactivating the product of *tufB*, another gene for EF-Tu, suppresses all three types of nonsense mutations [24].

Despite structural similarity on the polypeptide level, EF-1α cannot functionally substitute the *sup*2 product since earlier [4] a number of conditionally lethal mutants of *sup*2 were isolated. These data indicate that the *sup*2 protein is indispensible for viability of the yeast cell. Another characteristic distinguishes the yeast EF-1α and the *sup*2 gene product, namely the codon usage. EF-1α is one of the most abundant proteins in yeast and the codon usage in its gene is highly biased in good agreement with the results of Bennetzen and Hall [25]. In contrast, the *sup*2 gene does not show a high level of codon bias (not shown) suggesting that it does not belong to the highly expressed gene group.

Although a part of the *sup*2 gene product described in this paper possesses a high structural homology to EFs and in particular to yeast EF-1α and bacterial EF-Tu, the functional role of the *sup*2 gene product in protein synthesis seems to be different and remains to be established.

REFERENCES

[1] Inge-Vechtomov, S.G. (1964) Vestn. LGU (USSR) 9, 112–117.

[2] Inge-Vechtomov, S.G. and Andrianova, V.M. (1970) Genetika (USSR) 6, 103–115.

[3] Hawthorne, D.C. and Leupold, U. (1974) Curr. Topics Microbiol. Immunol. 64, 1–47.

[4] Surguchov, A.P., Smirnov, V.N., Ter-Avanesyan, M.D. and Inge-Vechtomov, S.G. (1984) in: Physicochemical Biology Review (Skulachev, V.P. ed.) vol.4, pp.147–205, Soviet Scientific Reviews, Moscow.

[5] Himmelfarb, H.J., Maicas, E. and Friesen, J.D. (1985) Mol. Cell. Biol. 5, 816–822.

[6] Surguchov, A.P., Berestetskaya, Yu.V., Fominykch, E.S., Pospelova, E.M., Smirnov, V.N., Ter-Avanesyan, M.D. and Inge-Vechtomov, S.G. (1980) FEBS Lett. 111, 175–178.

[7] Eustice, D.C., Wakem, L.P., Wilhelm, J.M. and Sherman, F. (1986) J. Mol. Biol. 188, 207–244.

[8] Breining, P. and Piepersberg, W. (1986) Nucleic Acids Res. 14, 5187–5197.

[9] Breining, P., Surguchov, A.P. and Piepersberg, W. (1984) Curr. Genet. 8, 467–470.

[10] Telckov, M.V., Surguchov, A.P., Dagckesamanskaya, A.R. and Ter-Avanesyan, M.D. (1986) Genetika (Russian) 22, 17–25.

[11] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.

[12] Harr, R., Fallman, P., Haggstrom, M., Wahlstrom, L. and Gustafsson, P. (1986) Nucleic Acids Res. 14, 273–284.

[13] Nagata, S., Nagashima, K., Tsunetsugu-Yokota, Y., Fujimara, K., Miyazaki, M. and Kaziro, Y. (1984) EMBO J. 3, 1825–1830.

[14] Shirmaier, F. and Philippsen, P. (1984) EMBO J. 3, 3311–3315.

[15] Von Hemmert, F.J., Amons, R., Pluijms, W.J.M., Van Ormondt, H. and Moller, W. (1984) EMBO J. 3, 1109–1113.

[16] Nagata, S., Tsunetsugu-Yokota, Y., Naito, A. and Kaziro, Y. (1983) Proc. Natl. Acad. Sci. USA 80, 6192–6196.

[17] Yokota, T., Sugisaki, H., Takanami, M. and Kaziro, Y. (1980) Gene 12, 25–31.

[18] Dayhoff, M.O., Schwartz, R.M. and Orourt, B.C. (1978) in: Atlas of Protein Sequence and Structure 5, Suppl.3, 345–352.

[19] Kohno, K., Vchida, T., Ohkubo, H., Nakanishi, S., Nakashi, T., Fuku, T., Ohtsuka, E., Ikehara, M. and Okada, Y. (1986) Proc. Natl. Acad. Sci. USA 83, 4978–4982.

[20] Laursen, R.A., L'Italien, J.J., Nugarkatti, S. and Miller, D.L. (1981) J. Biol. Chem. 256, 8102–8109.

[21] Arai, K., Kawakita, M., Nakamura, S., Ishikava, I. and Kaziro, Y. (1974) J. Biochem. (Tokyo) 76, 523–534.

[22] Nakamura, S. and Kaziro, Y. (1983) J. Biochem. (Tokyo) 90, 1117–1124.

[23] Tapio, S. and Kurland, C.G. (1986) Mol. Gen. Genet. 205, 186–188.

[24] Vijgenboom, E., Vink, T., Kraal, B. and Bosch, L. (1985) EMBO J. 4, 1049–1052.

[25] Bennetzen, J.L. and Hall, B.D. (1982) J. Biol. Chem. 257, 3026–3031.

Gene, 66 (1988) 45–54
Elsevier

GEN 02398

Nucleotide sequence of the *SUP2* (*SUP35*) gene of *Saccharomyces cerevisiae*

(Omnipotent suppressor; translation ambiguity; gene structure; codon bias analysis; gene homology; elongation factor EF-1α; intron; mitochondrial import)

Vitaliy V. Kushnirov, Michail D. Ter-Avanesyan, Miroslav V. Telckov, Andrei P. Surguchov, Vladimir N. Smirnov and Sergey G. Inge-Vechtomov*

*Institute of Experimental Cardiology, U.S.S.R. Cardiology Research Center, Moscow (U.S.S.R.), and * Department of Genetics, Leningrad State University, Leningrad (U.S.S.R.) Tel. 218-1590*

Received 10 November 1987
Accepted 12 November 1987
Received by publisher 25 February 1988

SUMMARY

A nucleotide sequence of the yeast *Saccharomyces cerevisiae* omnipotent suppressor *SUP2* (*SUP35*) gene is presented. The sequence contains a single open reading frame (ORF) of 2055 bp, which may encode a 76.5-kDa protein. A single transcript of 2.3 kb corresponding to a complete ORF is found. Analysis of codon bias suggests that the *SUP2* gene is not highly expressed. The C-terminal part of the deduced amino acid sequence shows a high homology to yeast elongation factor EF-1α, whereas the N-terminal part is unique for the *SUP2* protein. The N terminus contains a number of short repeating elements and possesses an unusual amino acid composition.

Analysis of the nucleotide and deduced amino acid sequences indicates that three additional proteins could possibly be expressed, two of which might be initiated on internal ATG codons and a third might be formed by alternative splicing. One of these proteins is supposed to be imported into mitochondria. Possible functions of the *SUP2* gene product(s), especially its putative activity as a soluble factor controlling the fidelity of translation, are discussed.

INTRODUCTION

Studies of informational suppression have proved to be useful in elucidating the mechanisms of control of translational fidelity. In all cases studied so far informational suppression results from mutational alterations in the components of protein synthesis apparatus — usually either tRNAs, or ribosomal constituents (Ozeki et al., 1980; Sherman, 1982; Dequard-Chablat et al., 1986; Steege and Söll,

Correspondence to: Dr. V.V. Kushnirov, Institute of Experimental Cardiology, U.S.S.R. Cardiology Research Center, 3rd Cherepkovskaya Street 15a, 121552 Moscow (U.S.S.R.) Tel. 149-0559.

Abbreviations: aa, amino acid(s); bp, base pair(s); EF, elongation factor; kb, kilobases or 1000 bp; MBN, mung-bean nuclease; MBN buffer, see MATERIALS AND METHODS, section b; nt, nucleotide(s); ORF, open reading frame; Pipes, piperazine-*N*,*N*′-bis[2-ethanesulfonic acid]; Pollk, Klenow (large) fragment of *E. coli* DNA polymerase I; S1 buffer, see MATERIALS AND METHODS, section b; tRNA, transfer RNA; u, unit(s).

1979). Recently, nonsense-suppressor mutations in *tuf* genes coding for EF-Tu have been described in *Escherichia coli* (Vijgenboom et al., 1985). However, similar mutations in eukaryotes have not yet been reported.

For the past several years we were studying recessive omnipotent suppressors in yeast. It was shown that mutations in the genes named *SUP1* (*SUP45*) and *SUP2* (*SUP35*) give rise to a variety of pleiotropic effects, including temperature sensitivity, drug sensitivity and respiratory deficiency. From these observations we concluded that the suppressor genes are essential for viability. Biochemical analysis indicates that suppressor mutations decrease the accuracy of translation and affect protein synthesis both in the cytoplasm and in mitochondria (Surguchov et al., 1984).

Recently, both the *SUP1* and *SUP2* genes were cloned (Breining et al., 1984; Telckov et al., 1986) and the nucleotide sequence of the *SUP1* gene was determined (Breining and Piepersberg, 1986). In this paper we report the nucleotide sequence of the *SUP2* gene. Part of the sequence shows significant homology to yeast EF-1α, suggesting that the *SUP2* gene product is not a canonical ribosomal protein, but rather a soluble translation factor. This essential protein appears to be present in minor quantities and probably has not been detected by biochemical methods. Further characterization of its role may reveal new essential features of the eukaryotic translation machinery.

MATERIALS AND METHODS

(a) Subcloning and sequencing

A shuttle plasmid pSTR4 containing *SUP2* gene (Telckov et al., 1986), was used for the sequence determination. A set of subclones of *SUP2* gene sufficient for sequencing was obtained in two steps: (*i*) restriction fragments of *SUP2* were cloned into M13mp phages (M13mp10, 11, 18, 19), and (*ii*) in some cases subclones were further deleted using DNase I, as described (Lin et al., 1984). DNA restriction, ligation and other enzymatic treatments were carried out according to the suppliers' specifications (Pharmacia P-L Biochemicals). Transformation of *E. coli* (strain JM101) by M13 phage and purification of recombinant phage were done according to Messing (1983). The nucleotide sequence was determined using the dideoxy method of Sanger et al. (1977).

(b) Yeast RNA analysis

Preparation of total yeast RNA was performed, as described by Cottrelle et al. (1985). For the Northern analysis, 20 μg of total RNA were glyoxylated, electrophoresed on an agarose gel and transferred to nitrocellulose (Maniatis et al., 1982). RNA blots were hybridized with strand-specific M13 probes, which were prepared according to Messing (1983).

To obtain a single-stranded ^{32}P-labelled probe for 5′-end mapping, an M13 clone containing fragment *Kpn*I-*Bcn*I (bp 164 to −205, see Fig. 2) with the *Kpn*I site proximal to a sequencing primer site was used. The probe was synthesized by the primer extension with PolIk, cleaved at the 3′ end with *Bcn*I and separated from the template using a 5% polyacrylamide gel (Leer et al., 1984). Hybridization was performed for 6 h at 46°C in 80% formamide, 0.4 M NaCl, 0.4 M Pipes (pH 6.5) and 1 mM EDTA. Total yeast RNA (50 μg) and 100000 cpm of probe were used for each experiment. Then hybridization mixtures were diluted ten-fold with S1 or MBN buffer (30 mM Na · acetate, pH 4.6, 1 mM $ZnSO_4$, 250 mM (50 mM for MBN buffer) NaCl, 20 μg/ml of sonicated and denatured calf thymus DNA), supplemented with 1000–4000 u/ml of the appropriate nuclease and digested for 30 min at 37°C. After chloroform extraction the protected DNA fragments were precipitated with isopropanol in the presence of carrier tRNA and analyzed on a 5% polyacrylamide, 7 M urea sequencing gel.

RESULTS AND DISCUSSION

(a) Nucleotide sequence

For sequence analysis, a shuttle plasmid pSTR4 (Telckov et al., 1986) carrying a minimal fragment of the cloned yeast genomic DNA complementing temperature-sensitive *sup2* mutation was used. The restriction map for this fragment and sequencing

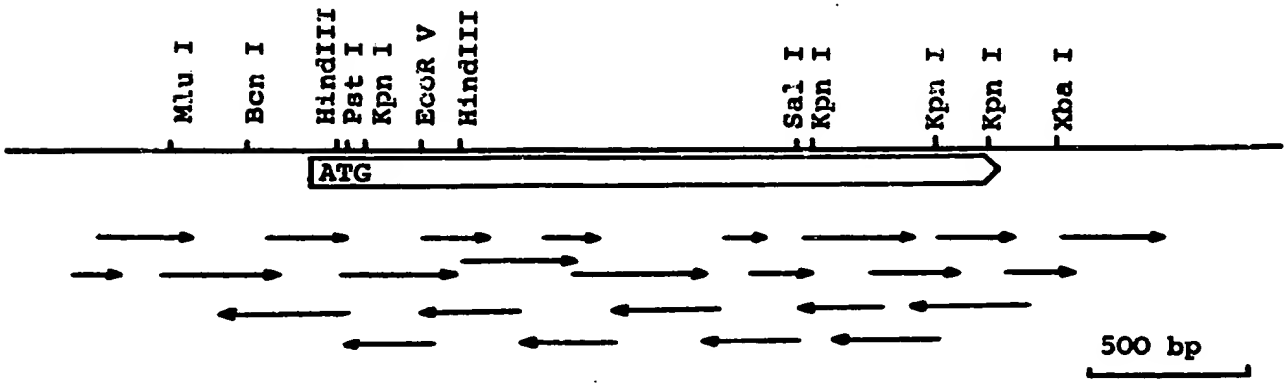


Fig. 1. Restriction map and sequencing strategy for the *SUP2* gene. Position and orientation of the ORF is represented by the arrowed open bar. Small arrows indicate the direction and extent of sequence determination on individual clones.

strategy are shown in Fig. 1. A contiguous sequence of 3320 nt containing a single long ORF of 2055 nt capable of coding for a protein of 76 545 Da was determined (Fig. 2).

(b) mRNA analysis

Hybridization of total yeast RNA with a single-stranded M13 probe containing a *Pst*I-*Xba*I fragment revealed a single band of 2.3 kb. The opposite strand of the same fragment did not hybridize with RNA. By S1 and MBN mapping two major transcription start points were found at nt positions −15 and −37, as well as two minor sites at nt positions −57 and −43 (Fig. 3), before the first ATG codon in the ORF. Taking into consideration that the size of the transcript is 2.3 kb, we conclude that the transcript contains the full length of the ORF. The first ATG in the ORF is present on the transcript and therefore it is the most probable initiator of trans-

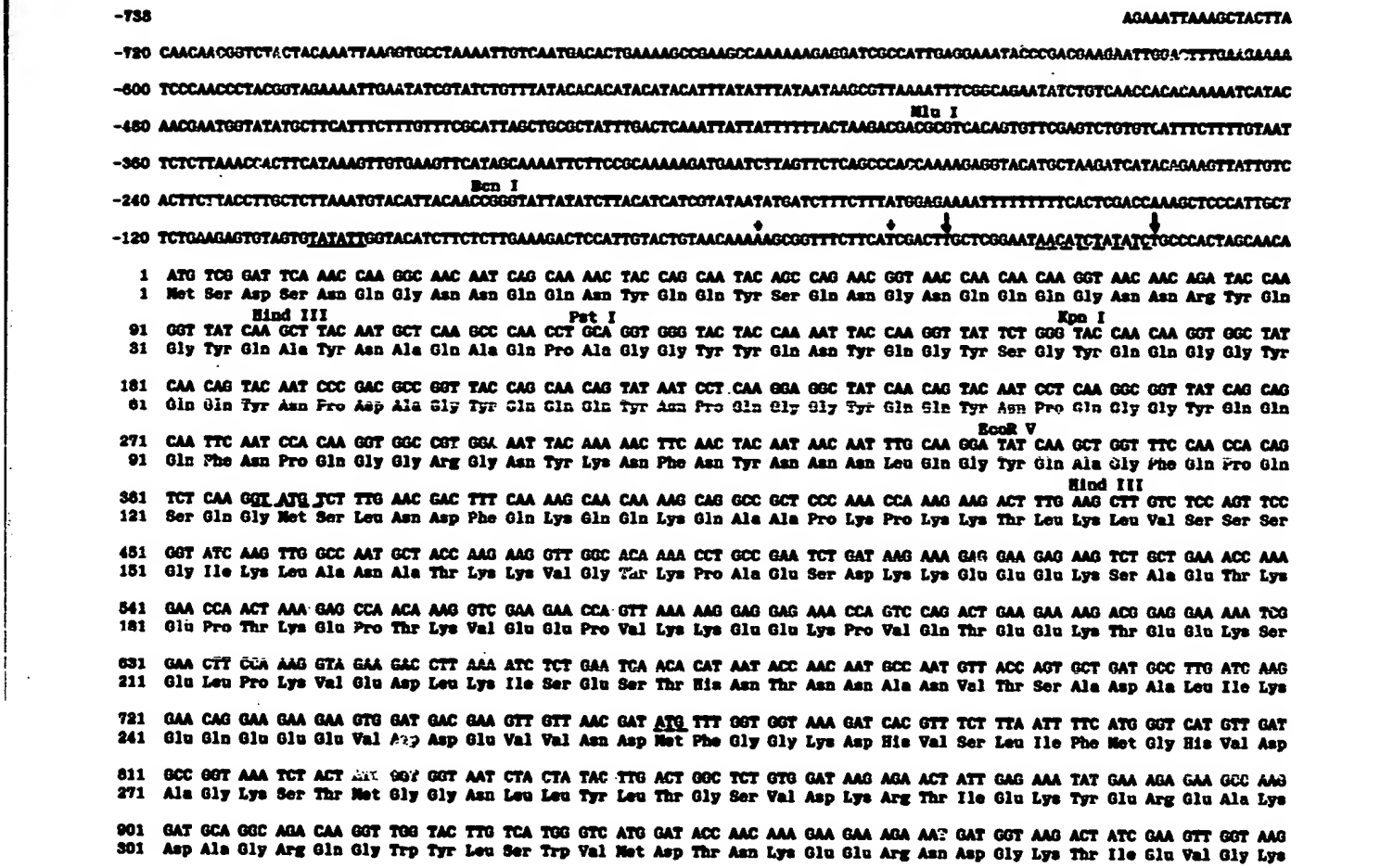
TABLE I

Codon usage in the *SUP2* gene

aa	Codon[a]		aa	Codon		aa	Codon		aa	Codon	
Phe	TTT	9	Ser	TCT	12	Tyr	TAT	10	Cys	TGT	4
Phe	TTC	7	Ser	TCC	7	Tyr	TAC	25	Cys	TGC	1
Leu	TTA	7	Ser	TCA	6	*ter*[b]	TAA	1	*ter*	TGA	0
Leu	TTG	15	Ser	TCG	3	*ter*	TAG	0	Trp	TGG	4
Leu	CTT	3	Pro	CCT	10	His	CAT	7	Arg	CGT	6
Leu	CTC	0	Pro	CCC	2	His	CAC	6	Arg	CGC	0
Leu	CTA	7	Pro	CCA	18	Gln	CAA	40	Arg	CGA	0
Leu	CTG	3	Pro	CCG	0	Gln	CAG	13	Arg	CGG	0
Ile	ATT	17	Thr	ACT	14	Asn	AAT	24	Ser	AGT	5
Ile	ATC	12	Thr	ACC	16	Asn	AAC	21	Ser	AGC	2
Ile	ATA	3	Thr	ACA	8	Lys	AAA	28	Arg	AGA	11
Met	ATG	19	Thr	ACG	1	Lys	AAG	38	Arg	AGG	1
Val	GTT	26	Ala	GCT	20	Asp	GAT	21	Gly	GGT	45
Val	GTC	10	Ala	GCC	16	Asp	GAC	9	Gly	GGC	10
Val	GTA	9	Ala	GCA	7	Glu	GAA	44	Gly	GGA	3
Val	GTG	5	Ala	GCG	0	Glu	GAG	13	Gly	GGG	2

[a] Codons, preferred in highly expressed yeast genes, are underlined.

[b] Symbol *ter* represents translational stop codons.



lation. It is interesting to note that sequences surrounding the two major transcription start points are similar to each other (Fig. 4).

(c) 5′- and 3′-flanking regions

A promoter element TATATT is located in a position typical for such elements in yeast, i.e., between bp −105 and −98 before the first ATG. The sequence AATAAA, which is thought to be a eukaryotic polyadenylation signal (Fitzgerald and Shenk, 1981), is situated 84–89 bp downstream from the terminating TAA. The sequence TAG...TAGT...TTT, a potential transcription termination signal in yeast (Zaret and Sherman, 1982), was found 115–142 bp downstream from the termination codon TAA. An interesting feature of the 3′-flanking region is the presence of the repeats $(TA)_{11}$ (95 bp downstream from TAA) and $(CAT)_{11}$ (350 bp downstream from TAA).

(d) Codon usage

The *SUP2* gene differs markedly in codon usage from highly expressed yeast genes, particularly ribosomal protein genes (Table I). The codon bias index according to Bennetzen and Hall (1982) was determined to be 0.42, whereas the range of values for ribosomal proteins is 0.79–0.94 (Sharp et al., 1986). Such a difference could mean, according to estimates given by Bennetzen and Hall (1982), that *SUP2* mRNA is at least an order of magnitude less abundant than mRNAs of ribosomal protein genes.

(e) Deduced amino acid sequence

A part of the amino acid sequence beginning with Met-254 (the third methionine in the sequence) is homologous to the full length of yeast EF-1α (Kushnirov et al., 1987). The remaining N-terminal part can be divided near the second methionine into

```
 991  GCC TAC TTT GAA ACT GAA AAA AGG CGT TAT ACC ATA TTG GAT GCT CCT GGT CAT AAA ATG TAC GTT TCC GAG ATG ATC GGT GGT GCT TCT
 331  Ala Tyr Phe Glu Thr Glu Lys Arg Arg Tyr Thr Ile Leu Asp Ala Pro Gly His Lys Met Tyr Val Ser Glu Met Ile Gly Gly Ala Ser

1081  CAA GCT GAT GTT GGT GTT TTG GTC ATT TCC GCC AGA AAG GGT GAG TAC GAA ACC GGT TTT GAG AGA GGT GGT CAA ACT CGT GAA CAC GCC
 381  Gln Ala Asp Val Gly Val Leu Val Ile Ser Ala Arg Lys Gly Glu Tyr Glu Thr Gly Phe Glu Arg Gly Gly Gln Thr Arg Glu His Ala

1171  CTA TTG GCC AAG ACC CAA GGT GTT AAT AAG ATG GTT GTC GTC GTA AAT AAG ATG GAT GAC CCA ACC GTT AAC TGG TCT AAG GAA CGT TAC
 391  Leu Leu Ala Lys Thr Gln Gly Val Asn Lys Met Val Val Val Val Asn Lys Met Asp Asp Pro Thr Val Asn Trp Ser Lys Glu Arg Tyr

1261  GAC CAA TGT GTG AGT AAT GTC AGC AAT TTC TTG AGA GCA ATT GGT TAC AAC ATT AAG ACA GAC GTT GTA TTT ATG CCA GTA TCC GGC TAC
 421  Asp Gln Cys Val Ser Asn Val Ser Asn Phe Leu Arg Ala Ile Gly Tyr Asn Ile Lys Thr Asp Val Val Phe Met Pro Val Ser Gly Tyr

1351  AGT GGT GCA AAT TTG AAA GAT CAC GTA GAT CCA AAA GAA TGC CCA TGG TAC ACC GGC CCA ACT CTG TTA GAA TAT CTG GAT ACA ATG AAC
 451  Ser Gly Ala Asn Leu Lys Asp His Val Asp Pro Lys Glu Cys Pro Trp Tyr Thr Gly Pro Thr Leu Leu Glu Tyr Leu Asp Thr Met Asn
          SalI                                                                                  Kpn I
1441  CAC GTC GAC CGT CAC ATC AAT GCT CCA TTC ATG TTG CCT ATT GCC GCT AAG ATG AAG GAT CTA GGT ACC ATC GTT GAA GGT AAA ATT GAA
 481  His Val Asp Arg His Ile Asn Ala Pro Phe Met Leu Pro Ile Ala Ala Lys Met Lys Asp Leu Gly Thr Ile Val Glu Gly Lys Ile Glu

1531  TCC GGT CAT ATC AAA AAG GGT CAA TCC ACC CTA CTG ATG CCT AAC AAA ACC GCT GTG GAA ATT CAA AAT ATT TAC AAC GAA ACT GAA AAT
 511  Ser Gly His Ile Lys Lys Gly Gln Ser Thr Leu Leu Met Pro Asn Lys Thr Ala Val Glu Ile Gln Asn Ile Tyr Asn Glu Thr Glu Asn

1621  GAA GTT GAT ATG GCT ATG TGT GGT GAG CAA GTT AAA CTA AGA ATC AAA GGT GTT GAA GAA GAA GAC ATT TCA CCA GGT TTT GTA CTA ACA
 541  Glu Val Asp Met Ala Met Cys Gly Glu Gln Val Lys Leu Arg Ile Lys Gly Val Glu Glu Glu Asp Ile Ser Pro Gly Phe Val Leu Thr

1711  TCG CCA AAG AAC CCT ATC AAG AGT GTT ACC AAG TTT GTA GCT CAA ATT GCT ATT GTA CAA TTA AAA TCT ATC ATA GCA GCC GGT TTT TCA
 571  Ser Pro Lys Asn Pro Ile Lys Ser Val Thr Lys Phe Val Ala Gln Ile Ala Ile Val Glu Leu Lys Ser Ile Ile Ala Ala Gly Phe Ser
                                                                                            Kpn I
1801  TGT GTT ATG CAT GTT CAT ACA GCA ATT GAA GAG GTA CAT ATT GTT AAG TTA TTG CAC AAA TTA GAA AAG GGT ACC AAC CGT AAG TCA AAG
 601  Cys Val Met His Val His Thr Ala Ile Glu Glu Val His Ile Val Lys Leu Leu His Lys Leu Glu Lys Gly Thr Asn Arg Lys Ser Lys

1890  AAA CCA CCT GCT TTT GCT AAG AAG GGT ATG AAG GTC ATC GCT GTT TTA GAA ACT GAA GCT CCA GTT TGT GTG GAA ACT TAC CAA GAT TAC
 631  Lys Pro Pro Ala Phe Ala Lys Lys Gly Met Lys Val Ile Ala Val Leu Glu Thr Glu Ala Pro Val Cys Val Glu Thr Tyr Gln Asp Tyr
                                  Kpn I
1980  CCT CAA TTA GGT AGA TTC ACT TTG AGA GAT CAA GGT ACC ACA ATA GCA ATT GGT AAA ATT GTT AAA ATT GCC GAG TAA ATTTCTTGCAAACAT
 661  Pro Gln Leu Gly Arg Phe Thr Leu Arg Asp Gln Gly Thr Thr Ile Ala Ile Gly Lys Ile Val Lys Ile Ala Glu ...

2073  AAGTAAATGCAAACACAATAATACCGATCATAAAGCATTTTCTTCTATATTAAAAAACAAGGTTTAATAAAGCTGTTATATATATATATATATATAGACGTATAATTAGTTTAGTTCT
                                                                                  Xba I
2193  TTTTGTACCATATACCATAAACAAGGTAAACTTCACCTCTCAATATATCTAGAATTTCATAAAAATATCTAGCAAGGTTTCAACTCCTTCAATCACGTTTTCATCATAACCCTTCCCCGG

2313  CGTTATTTCAGAATGTGCAAAATCTATTAGTGACATGGAACTCAAAGAACCAGTTGTTTTTTTGTCCTTTGGTCCTTCGCTGCTTCCCTCGGCATCATCATCATCATCATCATTATC

2433  ATCATCGTCGTCATCATCGTCTATAAAATCATCTCGCATAAGTTTGTCAACATCATTTAGTAATTCCCATCGCTCCGGGTCTCCTTCGTAAATAAACAAAAGACTACTTGATATCATTCT

2553  AACTTCTTCTTCTAGCATAGTATTATAAAA
```

Fig. 2. Nucleotide sequence and deduced amino acid sequence of the *SUP2* gene. The location of restriction sites is indicated. Sequence elements TATATT, AATAAA and TAG...TAGT...TTT (Zaret and Sherman, 1982), which may be relevant for initiation or termination of the transcription, are underlined by solid lines. Major and minor transcription start points are marked by downward arrows. Underlined by dashed lines are: HOMOL1-like sequence, the second and third in-frame ATG codons and sequences GTATGT and TACTAAC typical for yeast introns. Second ATG and the GTATGT sequence do overlap.

two fragments, both having an unusual amino acid composition (Table II; Fig. 5).

Region A is a region of 123 aa, beginning at the first methionine and contains repeats of three sequence elements, which make up most of its length (Fig. 6). Sequence Gln-Gly-Gly-Tyr-Gln-(Gln)-Gln-Tyr-Asn-Pro is repeated about six times (Fig. 6b). This region is rich in Gln (28%), Gly (17%), Asn (16%) and Tyr (16%), all four amino acids making up 78%.

Region B is a region of aa 124–253 rich in charged amino acids, Lys (18%) and Glu (18%), which may be further subdivided into four stretches: (*1*) a stretch of aa 124–164 is positively charged and resembles the signal sequences for mitochondrial import (von Heijne, 1976; see RESULTS, section g, for details); (*2*) a stretch of aa 165–222 contains several repeats of tetrapeptides: Lys-Lys-Glu-Glu, Thr-Lys-Glu-Pro, Glu-Glu-Lys-Ser, Thr-Glu-Glu-Lys; (*3*) a stretch of aa 223–235 does not contain charged residues; (*4*) a stretch of aa 236–253 carries a negative charge (9 aa residues out of 18 are Asp or Glu).

It is interesting that region B contains 24 Lys, but does not contain Arg.

(f) Possible existence of additional *SUP2* gene products

RNA analysis reveals a single transcript for the *SUP2* gene containing a complete ORF. However, a detailed analysis of the nucleotide and deduced amino acid sequences points to the possibility of existence of shorter transcripts and corresponding

protein products. One may suggest a possibility of translation initiation on the second and third ATG codons as well as excision of the part of the coding sequence resulting from splicing, as it is shown in sections g–i, below.

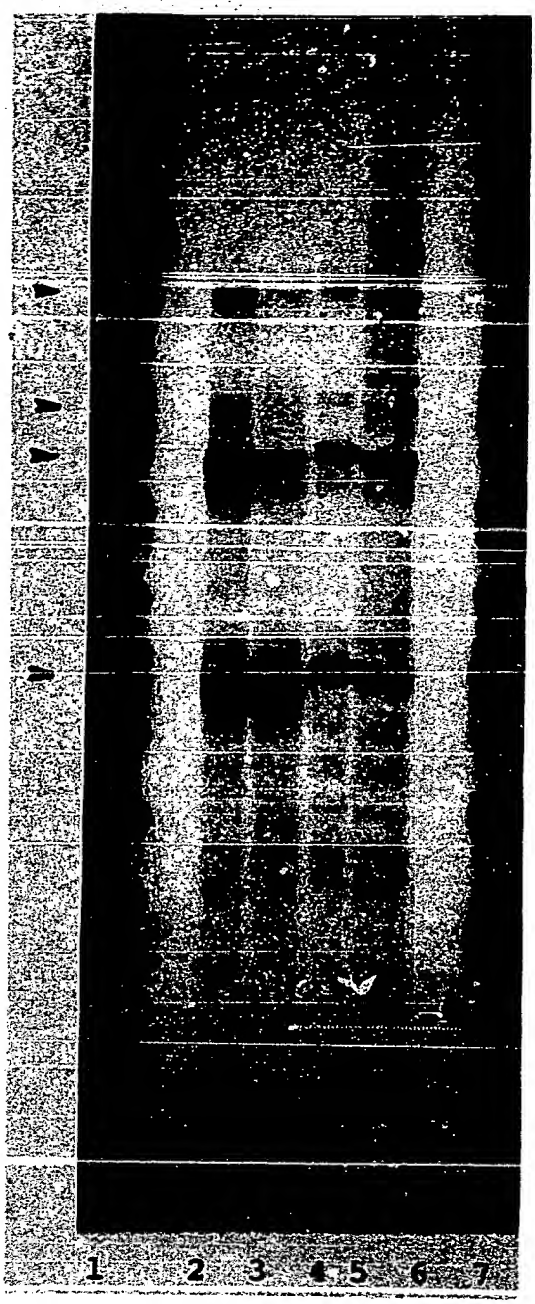


Fig. 3. Mapping of the 5′ end of the *SUP2* mRNA. Total yeast RNA was hybridized to a single-stranded ^{32}P-labeled *Kpn*I-*Bcn*I fragment (nt 165 to −205) at 46°C, in 80% formamide and treated with S1 or MB nuclease in the following concentrations: lanes: 2, 2000 u/ml of S1; 3, 4000 u/ml of S1; 4, 1000 u/ml of MBN; 5, 4000 u/ml of MBN; 6, control without yeast RNA, 2000 u/ml of S1. A dideoxy sequencing lane of a known sequence was used as a marker (lanes 1 and 7). The transcription start points are indicated by arrowheads.

(g) Initiation of the translation on the second in-frame ATG

As we have shown earlier, many *sup2* mutations cause a respiratory deficiency, reduction in mitochondrially synthesized cytochrome content and decrease in the rate of protein synthesis in mitochondria. These data allowed to predict the existence

```
                -37↓
                TCGACTTGCTCGGAA
Consensus:      T  AYYTGCYCR  A
                TATATCTGCCCACTA
                -15↑
```

Fig. 4. Similarity of the two major transcription start point regions. Transcription start points are marked by arrows. R designates purine (A or G), Y designates pyrimidine (T or C). The upstream sequence is on top, the downstream sequence is at the bottom (see Fig. 2), and the consensus sequence is in the middle line.

TABLE II

Amino acid composition of the *SUP2* gene[a]

aa	Region			The entire *SUP2* protein 1–685
	A 1–123	B 124–253	E 253–685	
Ala	6	9	28	43
Arg	2	0	16	18
Asn	20	7	18	45
Asp	2	7	21	30
Cys	0	0	5	5
Gln	35	6	12	53
Glu	0	23	34	57
Gly	21	2	37	60
His	0	1	12	13
Ile	0	3	29	32
Leu	1	7	27	35
Lys	1	24	41	66
Met	1	1	17	19
Phe	3	1	12	16
Pro	6	8	16	30
Ser	5	10	20	35
Thr	0	11	28	39
Trp	0	0	4	4
Tyr	20	0	15	35
Val	0	10	40	50
Total	123	130	432	685

[a] Amino acid composition of regions A (aa 1–123), B (aa 124–253), E (aa 254–685) and the entire *SUP2* protein is shown. Unusually high content of some amino acids is underlined.

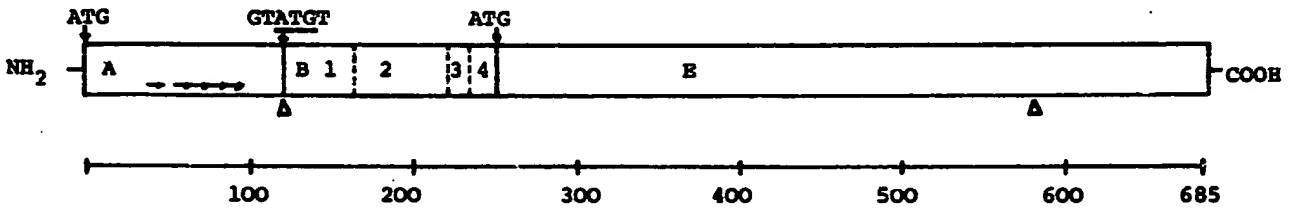


Fig. 5. Schematic representation of the predicted primary structure of the *SUP2* protein. Segment (A) represents a region containing extensive repeats of 8–10-bp sequence elements (shown by short horizontal arrows), and rich in Gln (28%), Gly (17%), Asn (16%) and Tyr (16%). Segment (B) represents a region with high content of charged residues Lys (18%) and Glu (18%). It may be subdivided into four stretches: (1) a positively charged stretch similar to signal sequences for mitochondrial import (von Heijne, 1986); (2) a stretch containing several tetrapeptide repeats rich in Lys and Glu; (3) a stretch without charged residues; (4) a stretch carrying a negative charge. Segment (E) represents a region homologous to the full length of yeast EF-1α. The sites corresponding to the first, second and third ATG codons are denoted by downward arrows. Putative splice sites are marked by triangles. The scale below the map is in aa. (See also Fig. 2 and Table II.)

of a *SUP2* gene product, which may be imported into mitochondria (Surguchov et al., 1984). Most of such imported proteins have a signal sequence at their N termini, which is positively charged and able to form an amphiphilic helix (von Heijne, 1986). A similar sequence element is present in a single site in the *SUP2* protein after the second methionine (aa 124–164).

Conserved 12-bp sequences, HOMOL1 and RPG, are present in the 5′-flanking regions of most of yeast ribosomal protein genes (Teem et al., 1984; Leer et al., 1985). The sequence HOMOL1 is also found in 5′-flanking regions of genes encoding EF-1α (Huet et al., 1985) and in the *SUP1* gene (Breining and Piepersberg, 1986). It has been proposed, that these sequences are required for the transcriptional regulation of the components of the translational apparatus (Huet et al., 1985). The HOMOL1 box is usually located before the TATA box at a distance of 150–400 bp upstream from the transcription start point. In the *SUP2* gene a sequence AACATC-TATATC similar to the HOMOL1 sequence, AACATC(T/C)(G/A)T(A/G)CA, is present. However, since this sequence is situated after presumed TATA box at nt positions −27 to −16 before the first ATG codon, it is possible that it regulates initiation of transcription at a site before the second ATG. A corresponding putative TATA box is located 150 bp upstream from the second ATG.

(h) Initiation of translation on the third in-frame ATG

Upon alignment of homologous regions of amino acid sequences of EF-1α and the *SUP2*, protein-initiating methionine of the EF-1α corresponds exactly to the third methionine in the *SUP2* protein, thus indicating possible involvement of the latter in the initiation of translation. The following observation confirms this suggestion. Upon deletion from the *SUP2* gene of a restriction fragment *Hin*dIII-*Hin*dIII (nt 99–434), the second in-frame ATG is removed and the reading frame beginning from the first ATG is disrupted. However, high copy number plasmids containing such deletions still complement certain temperature-sensitive *sup2* mutations. This result can be explained only by the existence of a protein initiated on the third in-frame ATG. The calculated M_r of this protein is 48 039.

A minor mRNA band of 1.4 kb, hybridizing to the coding segment of the *SUP2* gene, which has been observed by Surguchov et al. (1986), may correspond to a transcript initiated before the third ATG codon. However, such a band was not found in this study. A possible explanation for this discrepancy is that this transcript occurs in relatively small amounts depending on the conditions, used for growth.

(i) Possibility of alternative splicing

The contiguous ORF of the *SUP2* gene contains sequences that are typical for introns in *S. cerevisiae* genes, including a completely conserved sequence TACTAAC, which is present in all yeast introns. This sequence is found in the *SUP2* gene around bp 1700. At the 5′ end of yeast introns a sequence GTATGT or, less frequently, GTACGT is located. Both of these sequences are found in the ORF of the *SUP2* gene near bp 364 and 1046, respectively. A trinucleotide TAG (bp 1748) nearest to the sequence

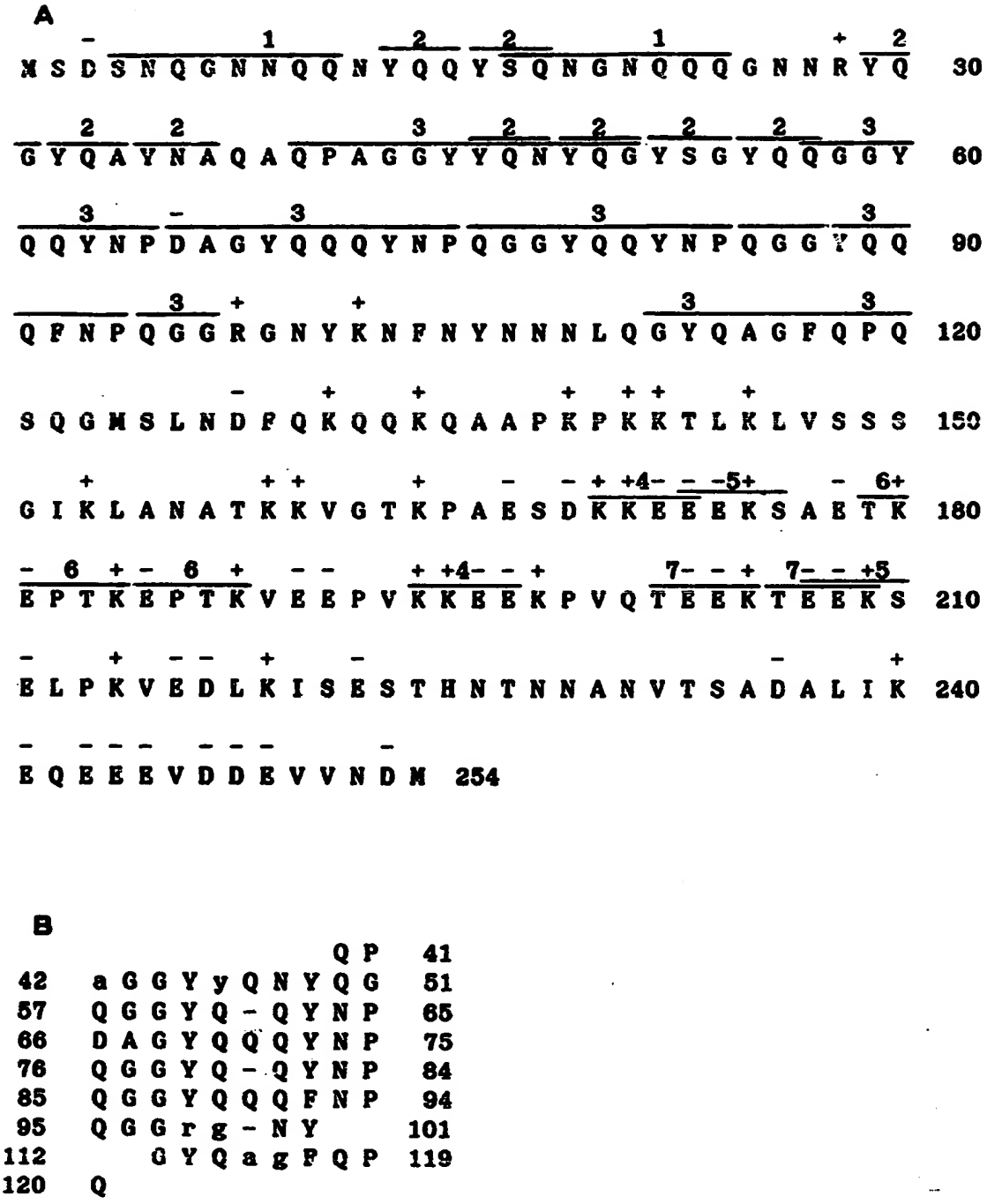


Fig. 6. Analysis of the amino acid sequence of the *SUP2* gene. (A) Deduced structure of the N-terminal region of the *SUP2* protein. The one-letter amino acid notation is used. Repeat sequences are overlined and numbered. Charged residues are marked with (+) or (−). (B) Alignment of the most extensive repeat element. One-aa gaps (dashes) were introduced in some places. Conservative amino acids are given in capitals.

TACTAAC may be regarded as a 3′ end of this hypothetical intron. The first of two donor splice sites (GTATGT) seems to be a more likely candidate for the 5′ end of the intron, because in this case the reading frame is not shifted by splicing. Location of this site is not random. The GTATGT sequence covers the second ATG codon and the border of the A and B regions of the deduced polypeptide (Fig. 5). The size of the protein product corresponding to spliced mRNA would be 25 kDa.

A large and functionally important part of the sequence lies inside the proposed intron, including

an assumed signal for mitochondrial import and domains, for which participation in GTP- and aminoacyl-tRNA-binding is predicted due to homology with EF-1α (Kushnirov et al., 1987). This indicates that the unspliced transcript must be expressed, even if splicing does occur. Such alternative splicing was not described in the yeast *S. cerevisiae*.

Although we did not detect multiple transcripts of the *SUP2* gene, their existence as well as expression of the corresponding protein products cannot be excluded.

(j) Possible functions of the *SUP2* protein

The C-terminal part of the *SUP2* protein, beginning from the third methionine (Met-254), shows significant homology to yeast EF-1α as well as to a family of analogous factors from other species (Kushnirov et al., 1987). The degree of amino acid homology amounts to 62%, considering conservative amino acid substitutions as homologous. Furthermore, nonhomologous stretches of significant length are absent in this region. This allows us to suggest that the *SUP2* protein possesses, apart from its N-terminal domains, the same functional domains as EF-1α, including GTP- and aminoacyl-tRNA-binding domains, where the degree of homology is highest. One might speculate then that these two proteins act at the same site on the ribosome and that their mode of action is rather similar. However, it is important to emphasize, that they are not interchangeable, since disruption of the *SUP2* gene is lethal (M.D.T.-A. and A.R. Dagckesamanskaya, in preparation). Furthermore, as pointed out in RESULTS, section d, the *SUP2* protein appears to be much less abundant, than EF-1α. Taken together, these and other data are consistent with the assumption that the product of the *SUP2* gene is a soluble factor that participates in the control of the fidelity of translation.

In a well-studied translation system of *E. coli*, a minor protein similar to the *SUP2* protein has not been found yet. At the same time, omnipotent suppressor mutations in EF-Tu have been described (Vijgenboom et al., 1985). Moreover, analysis of these mutants revealed a reduction in the accuracy of the protein synthesis at both the primary aminoacyl-tRNA selection and the proofreading steps (Tapio and Kurland, 1986). This allows us to suggest that EF-Tu, apart from a function analogous to EF-1α, may also perform a proofreading function, for which the *SUP2* protein is specialized in *S. cerevisiae*.

To determine the role of *SUP2* gene product(s) in protein synthesis it will be necessary to identify and purify the protein and study it biochemically. Genetic approaches and recombinant DNA techniques may also give valuable information, for example, examination of nucleotide substitutions, leading to suppression.

ACKNOWLEDGEMENTS

We are grateful to Dr. G.P. Samokhin and Dr. A.B. Sudarickov for the help with computer analysis of the data.

REFERENCES

Bennetzen, J.L. and Hall, B.D.: Codon selection in yeast. J. Biol. Chem. 257 (1982) 3026–3031.

Breining, P. and Piepersberg, W.: Yeast omnipotent suppressor *SUP1* (*SUP45*): nucleotide sequence of the wild-type and a mutant gene. Nucl. Acids Res. 14 (1986) 5187–5197.

Breining, P., Surguchov, A.P. and Piepersberg, W.: Cloning and identification of a DNA fragment coding for the *SUP1* gene of *Saccharomyces cerevisiae*. Curr. Genet. 8 (1984) 467–470.

Cottrelle, P., Thiele, D., Price, V.L., Memet, S., Micouin, J.Y., Marck, C., Buchler, J.M., Sentenac, A. and Fromageot, P.: Cloning, nucleotide sequence and expression of one of two genes, coding for yeast elongation factor 1α. J. Biol. Chem. 260 (1985) 3090–3096.

Dequard-Chablat, M., Coppin-Raynal, E., Picard-Bennoun, M. and Madjar, J.J.: At least seven ribosomal proteins are involved in the control of translational accuracy in a eukaryotic organism. J. Mol. Biol. 190 (1986) 167–175.

Fitzgerald, M. and Shenk, T.: The sequence 5'–AAUAAA–3' forms part of the recognition site for polyadenylation of late SV-40 messenger RNAs. Cell 24 (1981) 251–260.

Huet, J., Cottrelle, P., Cool, M., Vignais, M.-L., Thiele, D., Marck, C., Buchler, J.M., Sentenac, A. and Fromageot, P.: A general upstream binding factor for genes of the yeast translational apparatus. EMBO J. 4 (1985) 3539–3547.

Kushnirov, V.V., Ter-Avanesyan, M.D., Surguchov, A.P., Smirnov, V.N. and Inge-Vechtomov, S.G.: Localisation of possible functional domains in *SUP2* gene product of the yeast *Saccharomyces cerevisiae*. FEBS Lett. 215 (1987) 257–260.

Leer, R.G., Van Raamsdonk-Duin, M.M.C., Mager, W.H. and Planta, R.G.: The primary structure of the gene encoding yeast ribosomal protein L 16. FEBS Lett. 175 (1984) 371–376.

Leer, R.G., Van Raamsdonk-Duin, M.M.C., Mager, W.H. and Planta, R.G.: Conserved sequences upstream of yeast ribosomal protein genes. Curr. Genet. 9 (1985) 273–277.

Lin, H.C., Lei, S.P. and Wilcox, G.: An improved sequencing strategy. Anal. Biochem. 147 (1985) 114–119.

Maniatis, T., Fritsch, E.F. and Sambrook, J.: Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982, pp. 200–201.

Messing, J.: New M13 vectors for cloning. Methods Enzymol. 101 (1983) 20–78.

Ozeki, H., Iuokuchi, H., Yamao, F., Kodaira M., Sakano, H., Ikemura, T. and Shimura, Y.: Genetics of nonsense suppressor transfer RNAs in *Escherichia coli*. In Söll, D., Abelson, J.N., Schimmel, P.R. (Eds.), Transfer RNA: Biological Aspects. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1980, pp. 341–362.

Sanger, F., Nicklen, S. and Coulson, A.R.: DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74 (1977) 5463–5467.

Sharp, P.M., Tuohy, T.M.F. and Mosurski, K.R.: Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. Nucl. Acids Res. 14 (1986) 5125–5143.

Sherman, F.: Suppression in the yeast *Saccharomyces cerevisiae*. In Strathern, J.N., Jones, E.W. and Broach, J.R. (Eds.), Molecular Biology of the Yeast *Saccharomyces*: Metabolism and Gene Expression. Cold Spring Harbor Laboratory. Cold Spring Harbor, NY, 1982, pp. 463–486.

Steege, D.A. and Söll, D.G.: Suppression. In Goldberger, R.F. (Ed.), Biological Regulation and Development, Vol. 1. Plenum, New York, 1979, pp. 433–485.

Surguchov, A.P., Smirnov, V.N., Ter-Avanesyan, M.D. and Inge-Vechtomov, S.G.: Ribosomal suppression in eukaryotes. In Skulachev, V.P. (Ed.), Physicochemical Biology Reviews, Vol. 4. Soviet Scientific Reviews, Moscow, 1984, pp. 147–205.

Surguchov, A.P., Telckov, M.V. and Smirnov, V.N.: Absence of structural homology between *SUP1* and *SUP2* genes of yeast *Saccharomyces cerevisiae* and identification of their transcripts. FEBS Lett. 206 (1986) 147–150.

Tapio, S. and Kurland, C.G.: Mutant EF-Tu increases missense error in vitro. Mol. Gen. Genet. 205 (1986) 186–188.

Teem, J.L., Abovich, N., Kaufer, N.F., Schwindinger, W.F., Warner, J.R., Levy, A., Woolford, J., Leer, R.J., Van Raamsdonk-Duin, M.M.C., Mager, W.H., Planta, R.J., Schultz, L., Friesen, J.D., Fried, H. and Roshbash, M.: A comparison of yeast ribosomal protein gene DNA sequences. Nucl. Acids Res. 12 (1984) 8295–8312.

Telckov, M.V., Surguchov, A.P., Dagckesamanskaya, A.R. and Ter-Avanesyan, M.D.: Isolation of a chromosomal DNA fragment containing *SUP2* gene of the yeast *Saccharomyces cerevisiae*. Genetika 22 (1986) 17–25 (in Russian).

Vijgenboom, E., Vink, T., Kraal, B. and Bosch, L.: Mutants of elongation factor EF-Tu, a new class of nonsense suppressors. EMBO J. 4 (1985) 1049–1052.

von Heijne, G.: Mitochondrial targeting sequences may form amphiphilic helices. EMBO J. 5 (1986) 1335–1342.

Zaret, K.S. and Sherman, F.: DNA sequence required for efficient transcription termination in yeast. Cell 28 (1982) 563–573.

Communicated by A.A. Bayev.

The EMBO Journal vol.7 no.4 pp.1175–1182, 1988

A yeast gene required for the G_1-to-S transition encodes a protein containing an A-kinase target site and GTPase domain

Yoshiko Kikuchi, Hiroyuki Shimatake and Akihiko Kikuchi[1]

Department of Molecular Biology, Toho University, School of Medicine, 5-21-16 Omori-Nishi, Ohta-ku, Tokyo 143 and [1]Mitsubishi-Kasei Institute of Life Sciences, 11 Minamiooya, Machida-shi, Tokyo 194, Japan

Communicated by R.Monier

A new temperature-sensitive mutant of *Saccharomyces cerevisiae*, *gst1* (G_1-to-*S* transition) was isolated. At non-permissive temperature the mutant cells with large buds accumulated and DNA synthesis was substantially arrested. From the reciprocal experiment of temperature-shift and mating-factor treatment, it was shown that the execution point was post 'START'. This suggested that the mutation affected the G_1-to-S phase transition in the cell cycle. A DNA clone complementing the *gst1-1* mutation was isolated from a yeast gene library, and *gst1* was mapped in chr4R, by Southern blotting of cloned sequence to the individual yeast chromosome DNA by OFAGE system and by genetic analysis. The gene product was tentatively assigned from DNA sequencing analysis, as a protein of mol. wt 76 565 which contained consensus sequences for a target site of cAMP-dependent protein kinase(s) and for GTPase with extensive homology to polypeptide chain elongation factor EF1α.

Key words: cell cycle/DNA synthesis/EF1α homolog/site for A-kinase/yeast *ts* mutant

Introduction

To study the eukaryotic cell cycle, the budding yeast *Saccharomyces cerevisiae* offers a useful model system. Mutations causing stage-specific arrest in the cell cycle are important to clarify what kinds of gene products are involved in those complex processes, and the characterization of these proteins gives us clues to understand how they operate (Hartwell *et al.*, 1973; Pringle and Hartwell, 1981; Reed *et al.*, 1985; Patterson *et al.*, 1986). The initial event in the cell cycle, called 'START', takes place in the late G_1 phase, where the decision to undergo one cell cycle is made only when most of the cellular molecules are ready to be duplicated. Several *cdc* (*c*ell *d*ivision *c*ycle) mutants were isolated which affected this 'START' event (Hartwell, 1973; Hartwell *et al.*, 1974; Reed, 1980; Bedard *et al.*, 1981). Once a cell cycle has passed through 'START', DNA synthesis initiates and the normal sequential events of the cell cycle follow. For the subsequent step, but immediately prior to the DNA synthesis, two *cdc* genes, namely *cdc4* and *cdc7*, have been known to function in this order (Hereford and Hartwell, 1974). After the DNA is duplicated, each chromosome connected via *cis*-acting centromeres to the mitotic apparatus is faithfully distributed to each daughter cell. The yeast is the only organism whose centromere sequences have been cloned and characterized (Clarke and Carbon, 1980; Stinchcomb *et al.*, 1982).

Our initial aim was to isolate temperature-sensitive mutants in which proteins acting on centromeres were defective. For this purpose we took advantage of the fact that the chromosome copy number was rigorously regulated by the number of the centromere, although the mechanism which maintains the copy number of each chromosome to be one per haploid cell is yet unknown. When the centromere sequence is incorporated into multicopy plasmids such as yeast 2μ-plasmid, the copy number of the hybrid plasmid drops as low as one molecule per cell (Stinchcomb *et al.*, 1982). If proteins exist that control the chromosome copy number at the centromere level, we might expect to obtain mutants with thermolabile proteins where the hybrid 2 μ-plasmid with centromere sequence could amplify to a higher copy number at an intermediate temperature. The amplification might be facilitated through the intramolecular inversion system of the 2 μ-plasmid mediated by *FLP*-protein between two inverted repeat sequences (Futcher, 1986; Volkert and Broach, 1986). To detect amplified hybrid plasmids, we connected a defective *LEU2* (Beggs, 1978) to the plasmid as a selective marker which expressed only 5% of the wild-type level (Erhart and Hollenberg, 1983), so that we could get Leu^+ transformants when the copy number increased. Unexpectedly, we obtained a group of mutants in which the copy number of CEN-plasmid did not increase and whose cell cycles were arrested at the stage prior to DNA synthesis at high temperature. In this report, one of the mutants, *gst1*, was analysed. The new gene appeared to be required for the G_1-to-S phase transition and encoded a protein with a target site of cAMP-dependent protein kinase(s) (Cohen, 1985) and a GTPase domain with extensive homology to the polypeptide chain elongation factor EF1α (Nagata *et al.*, 1984; Schirmaier and Philippsen, 1984).

Results

Arrest phenotype of the gst1 mutant

A new temperature-sensitive (*ts*) mutant of *S.cerevisiae* was isolated as described in Materials and methods. The strains YK21-02 and YK21-03 were derived from a backcross of the original isolate. In the second cross of YK21-02 with a wild-type C5051-3D, the diploid was temperature resistant and the *ts* phenotype segregated 2+:2−, indicating that it was a single and recessive mutation (*gst1-1*). When asynchronous cultures of the strain YK21-02 or YK21-03 were grown at 26°C or 30°C (permissive temperature) and shifted to 36°C (non-permissive temperature), budded cells accumulated (74% of the total cells in the case of YK21-02 and 64% in YK21-03), as shown in Figure 1A. Although the size of the buds and the percentage of budded cells seemed to be variable by genetic backgrounds, the terminal

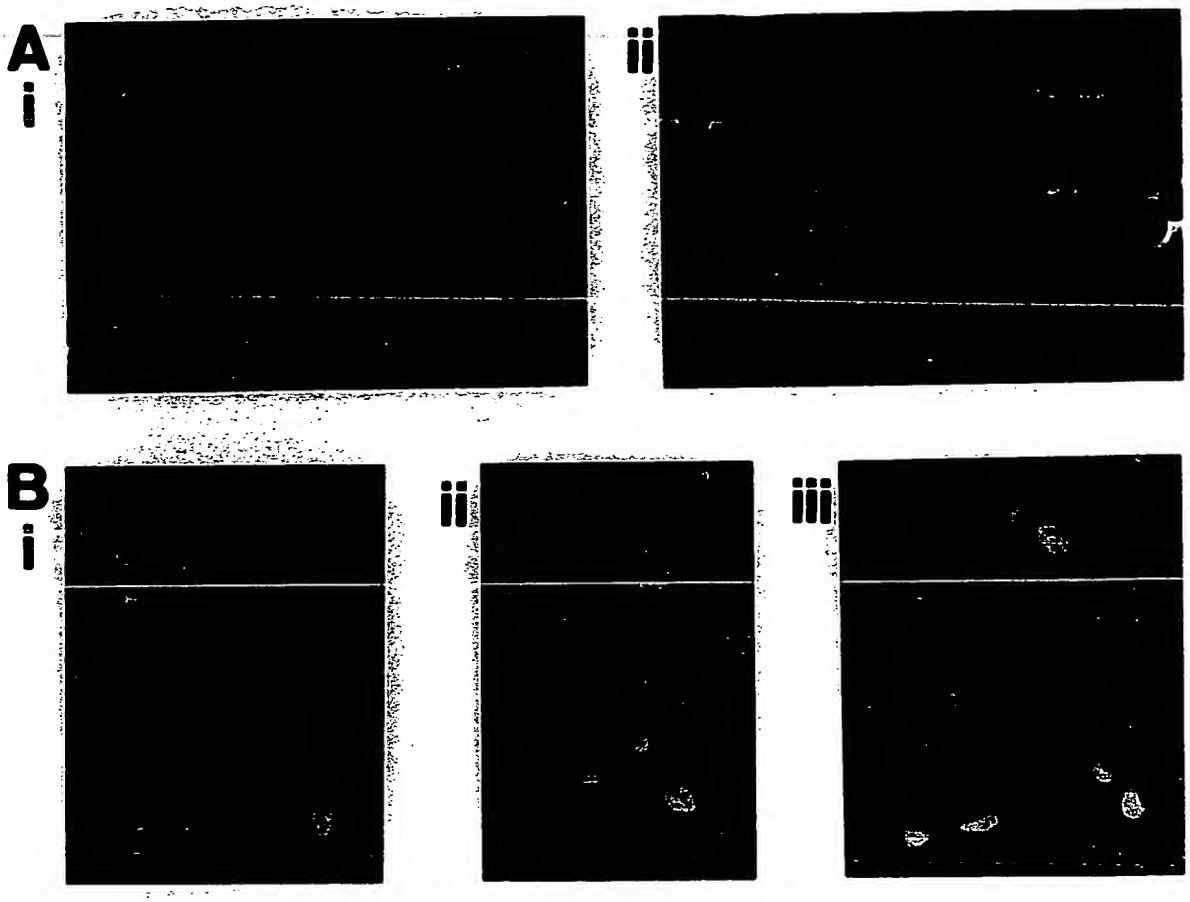


Fig. 1. Arrested morphology of the *gst1* mutant and location of the nuclei. (A) Cultures of the mutant cells were grown in YPD at 26°C and transferred to 36°C. At the following times after shift, photographs were taken: (i) *gst1-1* (strain YK21-02), 6 h; (ii) *gst1-1* (YK21-03), 6 h. (B) Mitochondrial DNA from yeast cells was removed as described in Materials and methods and nuclear DNAs were stained with DAPI. (i) Cells of *gst1-1* ϱ^0 (YK21-02 ϱ^0) incubated at 36°C for 6 h after shift. (ii) *gst1-1* ϱ^0 (YK21-02 ϱ^0), 36°C, overnight. (iii) *cdc7-1* ϱ^0 (124 ϱ^0), 36°C, 6 h. Bar = 5 μm.

morphology of the arrested cells appeared to be cell cycle specific.

In order to find at which stage of the cell cycle the mutant cells were blocked in the restrictive condition, we determined whether it was after the mating-pheromone-sensitive step or not. The α-mating factor arrests cells of *a* mating type in the late G_1 stage, before the initiation of DNA synthesis, and changes them to schmoo structures. Since the mating efficiency of the strain YK21-02 or YK21-03 was $<10\%$ (this phenotype was derived from *ste* of the original strain), we used the strain YK32-2C in which the phenotype of low mating efficiency segregated out in the second cross. After 3 h of treatment with α-factor, the cells were washed by filtration, transferred to a fresh medium (0 h, Figure 2Ai) and incubated at 37°C. Bud formation occurred and the buds grew bigger as the time increased (Figure 2Aii, iii). As a reciprocal experiment, arrested cells at the restrictive temperature (27°C, 3 h, Figure 2Bi) were mixed with α-factor and the culture was shifted to 26°C. Nuclear and cell divisions occurred and both mother and daughter cells changed to schmoos during 3 h (Figure 2Bii, iii). The cell number increased by 1.6-fold. Therefore, the execution point seemed to be after the mating-factor-sensitive step (post 'START').

DNA synthesis of mutant cells was measured as described in Materials and methods. Asynchronous cultures of the strain YK21-02 ϱ^0 missing its mitochondrial DNA were continuously labeled with [^{3}H]uracil at 26°C, and transferred to 36°C. After temperature shift, the DNA synthesis was substantially arrested, compared with the protein synthesis, which was followed by the incorporation of [^{35}S]methionine (Figure 3).

In the G_1-to-S phase transition of the cell cycle of *S.cerevisiae* two *cdc* genes, *cdc4* and *cdc7*, are well characterized. Temperature sensitivity of the strain *gst1* complemented either *cdc* mutant. The dumb-bell-shaped terminal morphology and defective DNA synthesis of *gst1* were similar to the phenotype of *cdc7*, except the location of nucleus. When the nuclear DNA of the strain YK21-02 ϱ^0 was fluorescently stained with DAPI, the nuclei appeared to locate near the junction of the buds (Figure 1B). In contrast, the arrested nucleus of *cdc7* migrated into the isthmus between the bud and the mother cell (Figure 1B; Hartwell *et al.*, 1973).

Isolation of plasmid capable of complementing gst1-1 mutation

A plasmid pYK801 capable of complementing the *gst1-1* mutation was isolated from gene libraries in multicopy vector YEp24 (Botstein *et al.*, 1979). The restriction map of the plasmid is shown in Figure 4. To localize the functional *GST1* gene, various DNA fragments were subcloned into a centromere vector YCp50 (Kuo and Campbell, 1983), which carried the yeast *URA3* gene, *ARS1* (*a*utonomously *r*eplicating *s*equence), *CEN4* (centromere 4) in pBR322 and

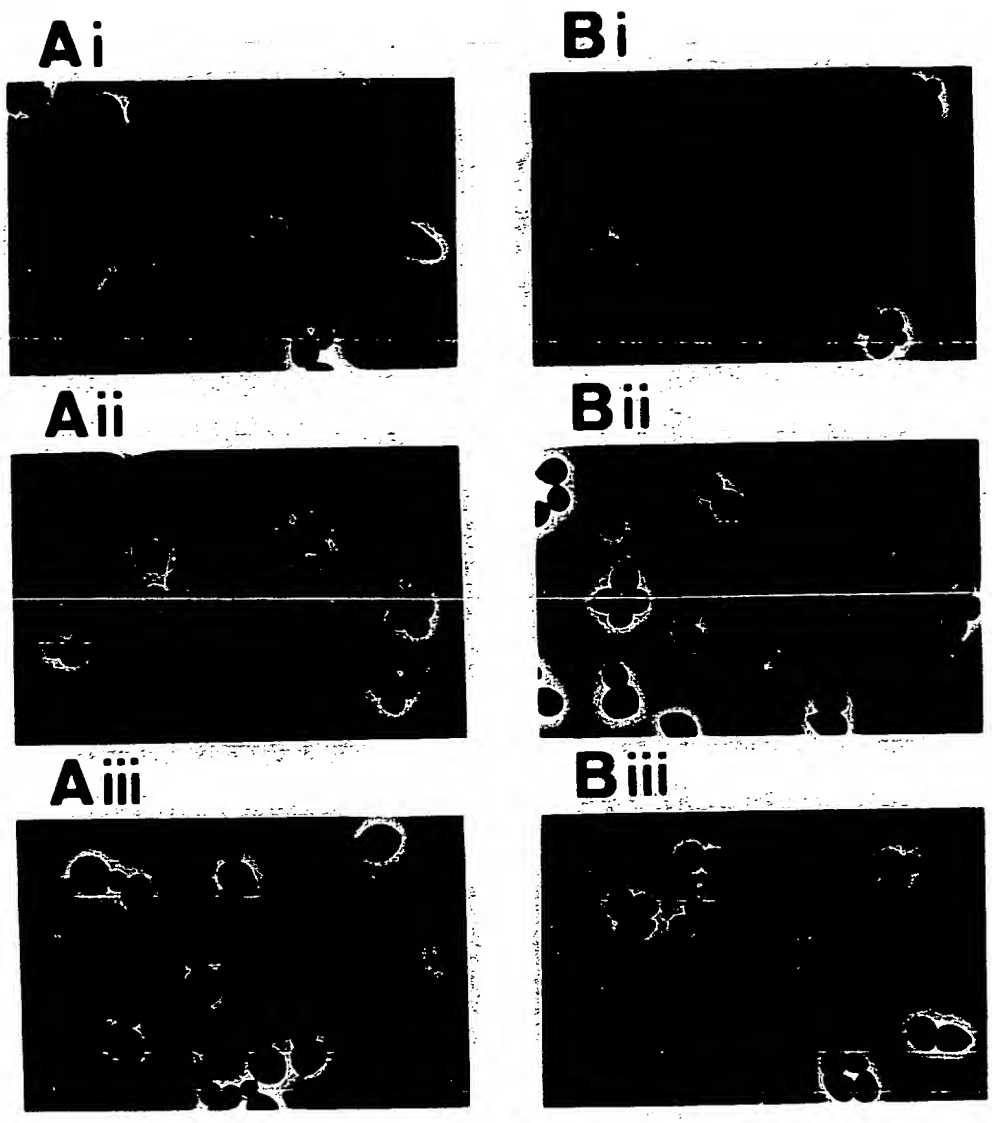


Fig. 2. Reciprocal experiment of temperature-shift and mating-factor treatment. (A) Cells of the strain YK32-2C were grown in YPD (pH 4) at 26°C and treated with 40 μg/ml α-factor for 3 h at 26°C. After washing α-factor by filtration, the culture was transferred to a fresh medium and grown further at 37°C. (i) 0 h, (ii) 2 h, (iii) 3 h after temperature shift. (B) Cells of YK32-2C were cultivated at 26°C and shifted to 37°C. After 3 h at restrictive temperature the cells were mixed with α-factor and transferred to 26°C to resume the growth: (i) 0 h, (ii) 2 h, (iii) 3 h incubation with α-factor.

behaved as a mini-chromosome in yeast. The resulting plasmids were introduced into YK21-02 and Ura$^+$ transformants were selected at 26°C. The temperature sensitivity of the transformants was checked by growth on a rich medium at 36°C. The complementing activity was localized within the 3.1-kb DNA region between the *Xba*I and *Pvu*II sites and the minimal requirement was the 1.3-kb region between the *Sal*I and the *Pst*I sites (Figure 4A), although the plasmid pYK825 carrying the 3.1-kb (*Eco*RI−*Pst*I) fragment partially complemented.

To confirm that the cloned DNA fragments contained the *GST1* gene itself and not an extragenic suppressor, the 4.6-kb (*Eco*RI−*Xho*I) fragment was inserted into an integration vector YIp5 (Scherer and Davis, 1979), which carried the *URA3* gene in pBR322. Since this hybrid, designated pYK821 (Figure 4B), cannot replicate autonomously in yeast, stable transformants arise only if the plasmid integrates into a chromosome by homologous recombination. To facilitate homologous recombination, the plasmid DNA was linearized with the restriction enzyme *Sal*I, the site of which was located in the middle of the gene and integrated into the chromosome (Orr-Weaver *et al.*, 1981). Ura$^+$ transformants were temperature-resistant recombinants. If they were mated with a wild-type strain (−+/+), only 5% of the spores turned out to be temperature sensitive upon meiosis, indicating that the cloned DNA was located at or close to the *gst1-1* mutation.

Mapping of the GST1 gene

To map the *GST1* gene, we performed the Southern blotting analysis to the yeast chromosomal DNAs fractionated in size by the orthogonal field-alternation gel electrophoresis (OFAGE) system (Carle and Olson, 1985), using pYK802 as a probe; the 8.9-kb *Eco*RI fragment of pYK801 carrying the *GST1* gene and pBR322 was self-ligated to make pYK802. The DNA probe hybridized to the top-most band to which the *TRP1* probe also hybridized, indicating that the cloned DNA was derived from chromosome 4 (data not shown).

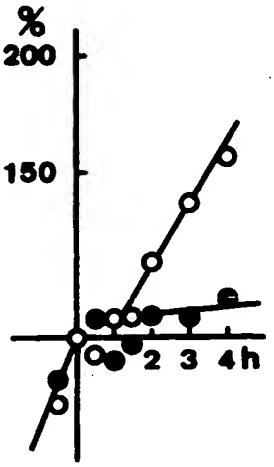


Fig. 3. DNA and protein synthesis. Assay procedure was described in Materials and methods. Results are expressed in percentages to the amount at the time of the temperature-shift. ○—○ protein; ●—● DNA.

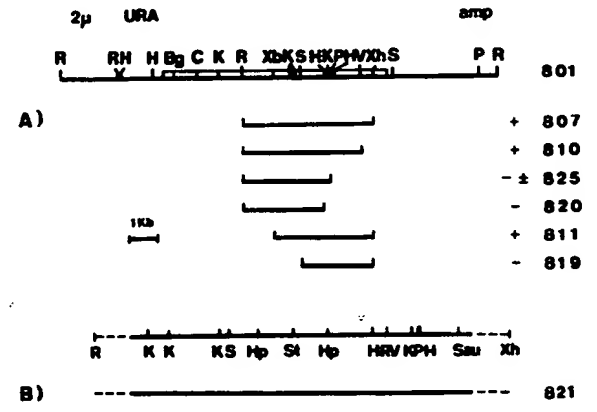


Fig. 4. Localization of the functional *GST1*. The restriction map of the plasmid pYK801 is drawn. YEp24 vector contains 2 μ-*ORI* and *URA3* on pBR322 (Botstein *et al.*, 1979). The yeast genomic DNA fragments have been cloned at the *Bam*HI site of the vector. The open box represents the insert. R, *Eco*RI; H, *Hin*dIII; Bg, *Bgl*II; C, *Cla*I; K, *Kpn*I; Xb, *Xba*I; S, *Sal*I; P, *Pst*I; V, *Pvu*II; Xh, *Xho*I.
(A) Complementation of *gst1* with various hybrid plasmids. DNA fragments from the insert of pYK801 were subcloned into YCp50 (Kuo and Campbell, 1983). Plasmid DNAs were introduced into *gst1* and Ura$^+$ transformants were selected at 26°C. Thermosensitivity of those transformants was checked at 36°C. pYK807: 4.6-kb (*Eco*RI−*Xho*I) fragment was cloned into YCp50. pYK810: DNA region between *Pvu*II site in the insert and *Nru*I site in the vector was deleted from pYK807. pYK825: 3.1-kb (*Eco*RI−*Pst*I) fragment and 2-kb (*Pst*I−*Hin*dIII) fragment of 2 μ-plasmid used as a linker, were inserted between *Eco*RI and *Hin*dIII sites of YCp50. pYK820: 2.8-kb (*Eco*RI−*Hin*dIII) fragment was cloned into YCp50. pYK811: 1.1-kb (*Eco*RI−*Xba*I) fragment in the insert of pYK807 was replaced with (*Eco*RI−*Xba*I) fragment containing *TRP1* region of YRp16 (Stinchcomb *et al.*, 1982). pYK819: 2-kb (*Eco*RI−*Sal*I) fragment of pYK807 was replaced with 2-kb (*Eco*RI−*Sal*I) fragment of YRp16.
(B) Plasmid structure for homologous integration. pYK821: 4.6-kb (*Eco*RI−*Xho*I) fragment was inserted between *Eco*RI and *Sal*I sites of YIp5 (Scherer and Davis, 1979). DNA sequence was determined (see Figure 6) in the region indicated by thick line. Hp, *Hpa*I; St, *Stu*I; RV, *Eco*RV; Sau, *Sau*3A.

Genetic evidence indicated that the *GST1* gene was located near the *CDC37* gene. In a cross of *gst1* with *cdc37*, the diploid was temperature resistant and after meiosis one out of 20 tetrads had a *ts*$^+$ recombinant, while other markers,

Fig. 5. Identification of *GST1* RNA transcript. A 10-μg sample of total RNA from strain DBY747*a* (1) or YNN27α (2) was denatured with glyoxal and resolved by 1.1% agarose gel electrophoresis, as described in Materials and methods. RNAs visualized by ethidium bromide staining were marked as size markers.

Trp and His segregated 2+:2−. Other *ts* mutants, *sec1*, *sec5* and *sec7*, mapped near *cdc37*, complemented with *gst1*. Therefore, the *GST1* gene may be a new *ts* gene near the *CDC37* gene on chromosome 4R.

RNA transcript

Total RNA was prepared from the wild-type strains DBY747*a* and YNN27α and a Northern blotting analysis was performed using the pYK812 DNA (pTZ18R containing the 1.3-kb *Kpn*I fragment from pYK801) as a probe. A major band was seen at 2.4 kb and a minor band appeared near 18S rRNA (Figure 5).

The nucleotide sequence of the GST1 locus

The nucleotide sequence of the cloned DNA fragment of the *GST1* gene and its flanking region was determined by dideoxy-chain termination method as shown in Figure 6. A predicted amino acid sequence (685 amino acids) for the long open reading frame (2055 bp) is also shown. The calculated mol. wt of the protein is 76 565. At several positions, including the sequence from −21 to −18 in the 5′ upstream region, there are TATA sequences. The nucleotide sequence downstream from the TAA termination codon contains sequences TAG···TAGT···TTT from +2170 to +2195, which may be associated with transcription termination and polyadenylation (Zaret and Sherman, 1982). A second potential *GST1* polyadenylation sequence, AATAAA is present at nucleotides +2139 to +2144 (Figure 6; Fitzgerald and Shenk, 1981; Bennetzen and Hall, 1982).

Homology with elongation factor 1α

The predicted *GST1* protein from the DNA sequence data is constituted from three domains. The domain I (codons 5−135) is rich in glutamine (30%), asparagine (16%) and tyrosine (15%), and highly conserved stretches of amino acids, QGGYQQ(Q)YNP, repeat about four times. In the

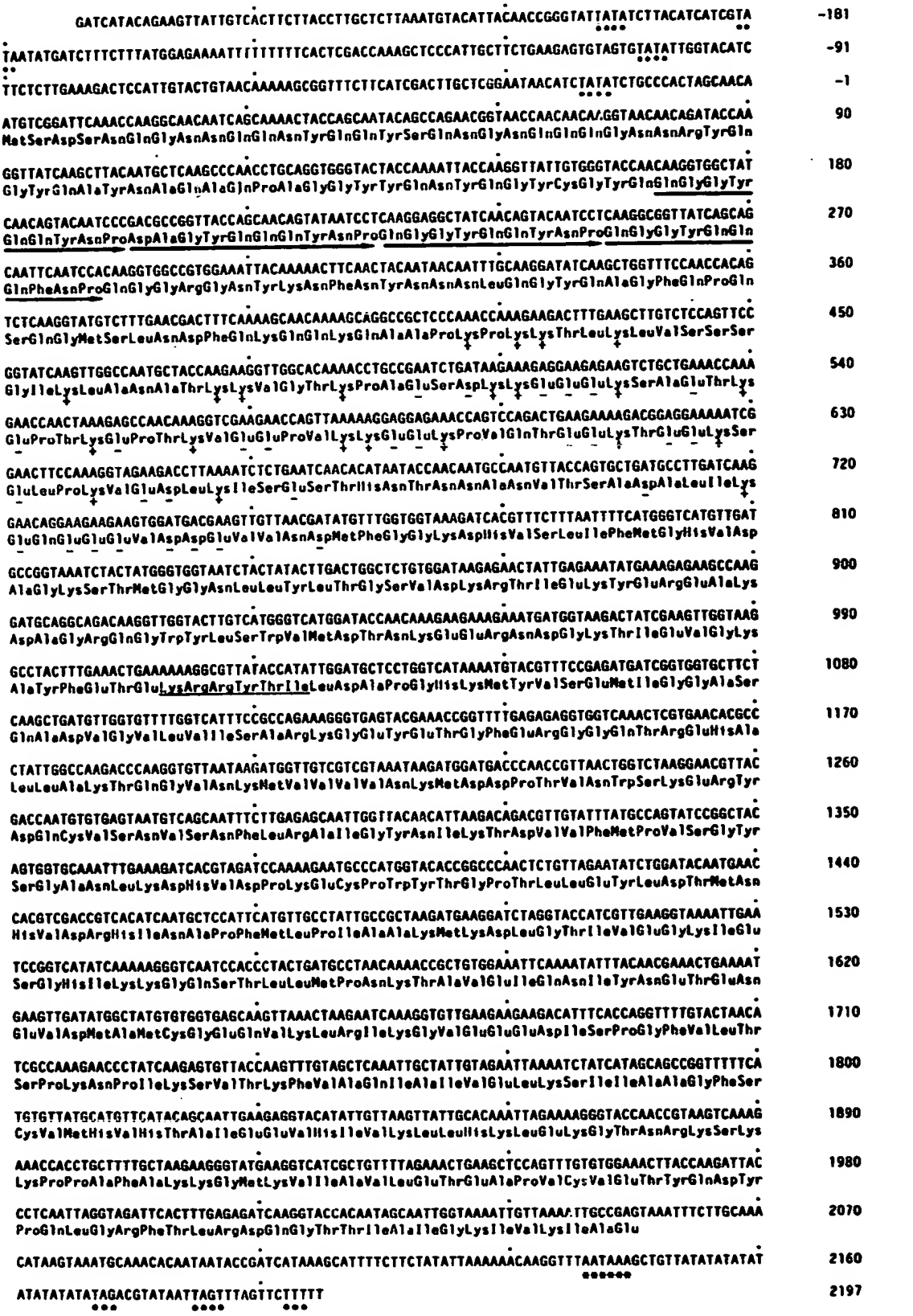


Fig. 6. Nucleotide sequence and predicted amino acid sequence of the *GST1* gene product. The coding sequence runs for 2055 nucleotides, which would encode a protein of 685 amino acids. The putative 'TATA' boxes are located at several positions marked by open circles. The potential signals for transcription termination and polyadenylation are indicated by closed circles and asterisks. The repetitive and highly conserved stretches of amino acids in domain I are shown by arrows. Basic or acidic amino acids in domain II are indicated as + or −. Comparison of amino acid sequence of domain III with EF1α is shown in Figure 7. The putative recognition site of A-kinase is underlined.

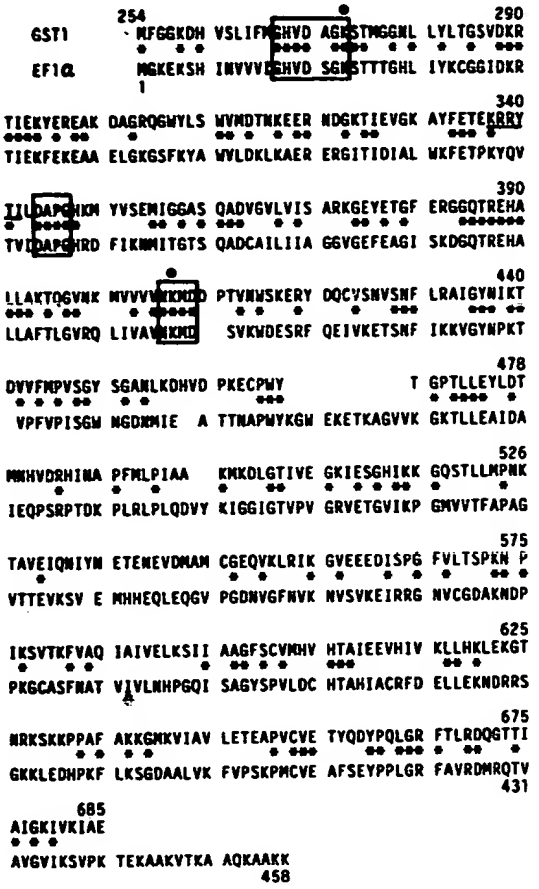


Fig. 7. Comparison of amino acid sequence of domain III of *GST1* with EF1α. Homologies between two sequences are indicated by asterisks. The region around 339 which might be recognized by A-kinase is underlined and the threonine at 341 might be phosphorylated. Three boxed regions specify consensus sequence elements in the GTP-binding domain (Dever *et al.*, 1987). Two lysine residues which might interact with GTP/GDP are marked by closed circles.

domain II (codons 139−249), acidic (especially Glu) and basic amino acids (mostly Lys) appear in turn as clusters. In the domain III which starts at codon 254 and ends at the carboxy terminus, extensive homology is found with yeast polypeptide chain elongation factor 1α (EF1α) (Nagata *et al.*, 1984; Schirmaier and Philippsen, 1984). The extent of homology with EF1α is 49% in the nucleotide sequence and 38% in the amino acid sequence (Figure 7). In particular, homology with the first half of EF1α is noteworthy; this is characteristic of a GTP-binding domain and GTPase activity center, as deduced from the X-ray crystallography of EF-Tu protein, an elongation factor of *Escherichia coli* (Jurnak, 1985). Three consensus sequence elements in the GTP-binding domain with distinct spacing (Dever *et al.*, 1987) are boxed in Figure 7. Two lysine residues at codons 273 and 407 might interact directly with GTP/GDP. Less homology is evident in the second half of EF1α, while a small part of the carboxy-terminal portion resumes some similarity between two sequences.

Furthermore, the sequence KRRYTI (codons 337−342) may represent a target site of cAMP-dependent protein kinase (A-kinase) (Cohen, 1985), which might phosphorylate the threonine residue at the position of codon 341.

Discussion

We have isolated many *ts* mutants using the selection described in Materials and methods. One group of these mutants accumulated budded cells after incubation at high temperature. Morphologically they are clearly different from 'START' mutants, which end up large unbudded cells. The new mutants proceed to the next stage and are arrested, like the *cdc7* mutant, with bud formation. The original selection could not be reproduced on the mutant described in this paper. It is not clear at this moment why we could collect such particular mutants. A CEN-plasmid was less stable at higher temperature and its copy number did not increase in the *gst1* mutant (unpublished observation).

The domain III of the *GST1* protein is strikingly similar to polypeptide chain elongation factor EF1α (Nagata *et al.*, 1984; Schirmaier and Philippsen, 1984). Since the stereochemical structure of the GTP-binding and GTPase domain of the EF-Tu protein of *E.coli* was already deduced from X-ray crystallography (Jurnak, 1985), we can superimpose the *GST1* protein on the EF-Tu protein with appropriate adjustments. Three consensus sequence elements for the GTP-binding domain in GTP-binding proteins are boxed in Figure 7 (Dever *et al.*, 1987). The region between the first and second consensus elements is known to interact with its effector. In this region the consensus sequence for the A-kinase target site is located. It is interesting to see whether *GST1* protein interacts with some factor(s) and if its binding is regulated by the phosphorylation.

In the signal transmission pathway, G-proteins, including *RAS* proteins localized at the membrane, modulate protein kinases. Subsequent phosphorylation of the target proteins is one of the critical events in the onset of S phase. Considering that *GST1* protein carries a potential target site for A-kinase, it would function epistatically to the kinase function. So, it is unlikely that the protein is located at the membrane like the authentic G-proteins. Also, the protein does not contain a Cys residue near the carboxy terminus, required for membrane localization (Powers *et al.*, 1986).

The amino acid sequence of *GST1* protein is not conclusive yet, because we only deduced this from the DNA sequencing data. We have not examined whether the gene contains intron sequences or not. However, the size of the 2.1-kb coding region, 0.1-kb 3′-untranslated region, 5′-untranslated region plus poly(A) chain appeared to be reasonable, compared with the size of the 2.4-kb mRNA as shown in Figure 5. Moreover the nucleotide sequence did not contain perfect consensus signals for splicing in yeast, GTATGT···TACTAAC···CAG (Schatz *et al.*, 1987). In the complementation experiment shown in Figure 4 one of the boundaries of the functional region should be beyond the restriction enzyme *Pst*I site and may be close to the boundary since the plasmid pYK825 complemented partially. In fact, the *Pst*I site is located close to the N terminus (codon 41) in the putative protein. We need further studies to clarify these undefined problems.

None the less, the fact that the tentative *GST1* protein, one of the essential genes for G_1-to-S transition, contains a potential A-kinase target site and GTPase domain would help in the analysis of the cell-cycle-specific events in eukaryotic cell proliferation. We are currently attempting to establish these findings biochemically and to see how they are regulated with respect to the cell cycle.

Materials and methods

Strains and genetic manipulations

E.coli JA221 (Beggs, 1978) was used for propagating plasmids and JM105 for M13 phage growth (Davis *et al.*, 1986). The strains of *S.cerevisiae* used in this study are the following: YK6-42, *a ste ade1 leu2 ura3 trp1 his3 cir*^{0}] (Kikuchi, 1983); C5674-3B, *α trp1 arg4 lys7*; YK21-02, *α gst1-1 ura3 trp1 his3* (YK6-42 *gst1-1* × C5674-3B); YK21-03, *a gst1-1 ura3 trp1 leu2 ade1 arg4* (YK6-42 *gst1-1* × C5674-3B); XMF2-28, *α sec1-1 trp1* (Nishizawa); XMF9-10, *α sec5-24 trp1* (Nishizawa); XMF3-6, *α sec7-1 trp1* (Nishizawa); DBY747, *a his3 ura3 leu2 trp1* (Matsui); YNN27, *α trp1 ura3 gal2* (Matsui); C5051-3D, *a his4 leu2 thr4 lys7*; YK32-2C, *a gst1-1 leu2 lys7 thr4 his* (YK21-02 × C5051-3D); SR672-1, *a cdc37-1 ura1 cyh2 gal2*; 124, *a cdc7-1*. The *cdc* mutants were derived from Yeast Genetic Stock Center at Berkeley.

Media, methods of mating, tetrad analysis and isolation of ϱ^0 strains were as described by Sherman *et al.* (1986). For α-factor treatment, the cells were grown in YPD (pH 4.0) and 40 μg/ml α-factor (Peptide Institute, Japan) were added at a cell density of 2×10^7/ml.

Plasmids and transformation

The YEp24 library was kindly provided by D.Botstein. Total DNA was partially digested with the restriction enzyme *Sau*3A and the fragments were inserted into the *Bam*HI site of YEp24 (Botstein *et al.*, 1979). Vectors used in this study were YCp50 (Kuo and Campbell, 1983), YIp5 (Scherer and Davis, 1979), YRp16 and YCp19 (Stinchcomb *et al.*, 1982), pJDB219 (Beggs, 1978) and pTZ18R (Pharmacia). Plasmid DNA was prepared as described previously (Kikuchi and Toh-e, 1986). Yeast transformation was performed by alkali ion method (Ito *et al.*, 1983). When linear DNA was used, tRNA (Sigma) was added as a carrier.

Isolation of gst1 mutant

The original aim of the mutant selection was to isolate *ts* mutants in which proteins acting on the centromere were defective, as described in the Introduction. For this selection procedure, the following plasmids were constructed. Plasmid pYK2068 was made by inserting the 2.4-kb *Bgl*II fragment containing *CEN4* of YCp19 (Stinchcomb *et al.*, 1982) into the *Bam*HI site of pJDB219. Plasmid pJDB219 harbored the entire yeast 2 μ-plasmid sequence (but the *FLP* gene was destroyed) on pMB9 vector along with a partially defective *LEU2* as a selective marker (Beggs, 1978). Since the 5′ upstream region of the *LEU2* was defective, the expression of Leu^+ phenotype was so low that transformants were Leu positive only when cells contained multicopies of this plasmid (Erhart and Hollenberg, 1983). Plasmid pYK2090 was constructed by inserting the same *Bgl*II fragment of YCp19 containing *CEN4* into pYK2029 which carried the *FLP* gene and *URA3* as a selective marker as described in Kikuchi (1983).

The yeast strain YK6-42 was mutagenized with ethyl methanesulfonate (8% survivals), grown in a minimal medium and cells were transformed with both plasmids pYK2068 and pYK2090. After 6 days of incubation at 34°C, ~8000 Ura^+ Leu^+ transformants were obtained. Most colonies were very tiny but 27 were relatively large. Fourteen out of 27 clones were *ts* mutants. One of these *ts* mutants was crossed with a wild-type strain C5674-3B, and YK21-02 and YK21-03 carrying the *gst1-1* allele were obtained.

Assays for macromolecular synthesis

The procedure for the measurement of DNA and protein synthesis was essentially as described by Johnston and Game (1978), except that the cells were continuously labeled with 20 μCi/ml [5,6-^{3}H]uracil (46 Ci/mmol; Amersham) or 6 μCi/ml [^{35}S]methionine (300 Ci/mmol; Amersham) in YPD medium, supplemented with 20 μg/ml adenine.

DAPI staining of nuclear DNA

Samples of 1 ml of ϱ^0 cells were removed at various times from a culture grown at 36°C (i.e. the non-permissive temperature). The cells were collected by centrifugation, washed once with 25% ethanol, 15 mM $MgCl_2$ and suspended in the same solution. They were allowed to fix at room temperature for at least 30 min and washed twice with cold water by centrifugation. The fixed cells were suspended in 0.5 μg/ml DAPI (4′,6′-diamidino-2-phenylindole, Sigma) in water and viewed through an Olympus DApo 100 UV lens on an Olympus BH2 microscope equipped for epifluorescence, and photographed with Kodax Tri-X film.

Orthogonal field-alternation gel electrophoresis (OFAGE)

The samples of yeast DNA were prepared by the embedded-agarose procedure described in Carle and Olson (1985). We designed and used a simplified model of the apparatus for OFAGE (Carle and Olson, 1984). The DNA samples were electrophoresed through 1.5% agarose gel (10 × 10 cm) in 45 mM Tris, 45 mM boric acid and 0.5 mM EDTA (pH 8.0), at 200 V with a switching interval of 50 s, for 20 h at 14°C. The gel was stained with 0.5 μg/ml ethidium bromide, extensively washed with deionized water and photographed on a UV transilluminator.

Southern and Northern analysis

Southern blotting analysis was performed as described previously (Kikuchi, 1983) except that the labeled probe was prepared with [^{32}P]dCTP using the Multiprime™ DNA labeling system (Amersham). For Northern analysis, total RNA was prepared as described by Maniatis *et al.* (1982). The RNA sample was incubated with 50% dimethylsulfoxide, 10 mM sodium phosphate (pH 7.0), 1 M glyoxal at 50°C for 60 min. Then the sample was fractionated by 1.1% agarose gel electrophoresis in 10 mM sodium phosphate (pH 7.0). Without any pretreatment, the gel was blotted to nitrocellulose filters. The bands of ribosomal RNAs were visualized by ethidium bromide staining and used as size markers.

DNA sequencing

Nucleotide sequences were determined by the dideoxy chain-termination method (Sanger *et al.*, 1977) using an M13-sequencing kit (Takara Shuzo Co., Kyoto). DNA fragments were cloned into M13 mp18 or mp19. Reaction products were resolved by electrophoresis through 8% acrylamide gels under denaturing conditions.

Acknowledgements

We would like to thank Drs D.Botstein, R.Davis, M.Nishizawa and Y.Matsui for providing plasmids or strains, and Yasushi Matsui for valuable discussions and help with some of the work in this study. We also thank Drs M.Ohba and Y.Oka for critical reading of the manuscript, Miharu Ando for technical assistance, Kumiko Yasui for computer search and Etsuko Ueno for preparing the manuscript.

References

Bedard,D.P., Johnston,G.C. and Singer,R.A. (1981) *Curr. Genet.*, **4**, 205–214.
Beggs,J.D. (1978) *Nature*, **275**, 104–108.
Bennetzen,J.L. and Hall,B.D. (1982) *J. Biol. Chem.*, **257**, 3018–3025.
Botstein,D.S., Falco,C., Stewart,S.E., Brennan,M., Scherer,S., Stinchcomb,D.T, Struhl,K. and Davis,R.W. (1979) *Gene*, **8**, 17–23.
Carle,G.F. and Olson,M.V. (1984) *Nucleic Acids Res.*, **12**, 5647–5664.
Carle,G.F. and Olson,M.V. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3756–3760.
Clarke,L. and Carbon,J. (1980) *Nature*, **287**, 504–509.
Cohen,P. (1985) *Eur. J. Biochem.*, **151**, 439–448.
Davis,L.G., Dibner,M.D. and Battey,J.F. (1986) *Basic Methods in Molecular Biology*. Elsevier, New York.
Dever,T.E., Glynias,M.J .and Merrick,W.C. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 1814–1818.
Erhart,E. and Hollenberg,C.P. (1983) *J. Bacteriol.*, **156**, 625–635.
Fitzgerald,M. and Shenk,T. (1981) *Cell*, **24**, 251–260.
Futcher,A.B. (1986) *J. Theor. Biol.*, **119**, 197–204.
Hartwell,L.H. (1973) *J. Bacteriol.*, **115**, 966–974.
Hartwell,L.H., Mortimer,R.K., Culotti,J. and Culotti,M. (1973) *Genetics*, **74**, 267–286.
Hartwell,L.H., Culotti,J., Pringle,J.R. and Reid,B.J. (1974) *Science*, **183**, 46–51.
Hereford,L.M. and Hartwell,L.H. (1974) *J. Mol. Biol.*, **84**, 445–461.
Ito,H., Fukuda,Y., Murata,K. and Kimura,A. (1983) *J. Bacteriol.*, **153**, 163–168.
Johnston,L.H. and Game,J.C. (1978) *Mol. Gen. Genet.*, **161**, 205–214.
Jurnak,F. (1985) *Science*, **230**, 32–36.
Kikuchi,Y. (1983) *Cell*, **35**, 487–493.
Kikuchi,Y. and Toh-e,A. (1986) *Mol. Cell. Biol.*, **6**, 4053–4059.
Kuo,C.-L. and Campbell,J.L. (1983) *Mol. Cell. Biol.*, **3**, 1730–1737.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Nagata,S., Nagashima,K., Tsunetsuga-Yokota,Y., Fujimura,K., Miyazaki,M. and Kaziro,Y. (1984) *EMBO J.*, **3**, 1825–1830.
Orr-Weaver,T.L., Szostak,J.W. and Rothstein,R.J. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 6354–6358.
Patterson,M., Sclafani,R.A., Fangman,W.L. and Rosamond,J. (1986) *Mol. Cell. Biol.*, **6**, 1590–1598.

Powers,S., Michaelis,S. Broek,D., Santa Anna-A,S., Field,J., Herskowitz,I. and Wigler,M. (1986) *Cell*, **47**, 413–422.
Pringle,J.R. and Hartwell,L.H. (1981) In Strathern,J.N., Jones,E.W. and Broach,J.R. (eds), *The Molecular Biology of the Yeast Saccharomyces*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 97–142.
Reed,S.I. (1980) *Genetics*, **95**, 561–577.
Reed,S.I., de Borros Lopes,M.A., Ferguson,J., Hadwiger,J.A., Ho,J.-Y., Horwitz,R., Jones,C.A., Lorincz,A.T., Mendenhall,M.D., Peterson,T.A., Richardson,S.L. and Wittenberg,C. (1985) *Cold Spring Harbor Symp. Quant. Biol.*, **50**, 627–634.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schatz,P.J., Pillus,L., Grisafi,P., Solomon,F. and Botstein,D. (1987) *Mol. Cell. Biol.*, **6**, 3711–3721.
Scherer,S. and Davis,R.W. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4951–4955.
Schirmaier,F. and Philippsen,P. (1984) *EMBO J.*, **3**, 3311–3315.
Sherman,F., Fink,G.R. and Hicks,J.B. (1986) *Laboratory Course Manual for Methods in Yeast Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Stinchcomb,D.T., Mann,C. and Davis,R.W. (1982) *J. Mol. Biol.*, **158**, 157–179.
Volkert,F.C. and Broach,J.R. (1986) *Cell*, **46**, 541–550.
Zaret,K.S. and Sherman,F. (1982) *Cell*, **28**, 563–573.

Received on July 15, 1987; revised on December 22, 1987

YEAST VOL. 6: 461–472 (1990)

Divergence and Conservation of *SUP2(SUP35)* Gene of Yeasts *Pichia pinus* and *Saccharomyces cerevisiae*

VITALY V. KUSHNIROV, MICHAEL D. TER-AVANESYAN‡, SVETLANA A. DIDICHENKO, VLADIMIR N. SMIRNOV, YURI O. CHERNOFF*, IRINA L. DERKACH*, OLGA N. NOVIKOVA*, SERGEY G. INGE-VECHTOMOV*, MICHAEL A. NEISTAT† AND ILYA I. TOLSTORUKOV†

Institute of Experimental Cardiology, U.S.S.R. Cardiology Research Center, 3rd Cherepkovskaya street 15a, 121552 Moscow, U.S.S.R.
**Department of Genetics, Leningrad State University, Universitetskaya nab. 7/9, 199034 Leningrad, U.S.S.R.*
†All-Union Institute for Genetics and Selection of Industrial Microorganisms, 1st Dorozhny proyezd 1, 113545 Moscow, U.S.S.R.

Received 7 December 1989; revised 27 March 1990

SUP2(SUP35) is an omnipotent suppressor gene, coding for an EF-1α-like protein factor, intimately involved in the control of translational accuracy in yeast *Saccharomyces cerevisiae*.

In the present study a *SUP2* gene analogue from yeast *Pichia pinus* was isolated by complementation of the temperature-sensitive *sup2* mutation of *S. cerevisiae*.

The nucleotide sequence of the *SUP2* gene of *P. pinus* codes for a protein of 82·4 kDa, exceeding the Sup2 protein of *S. cerevisiae* by 6 kDa. Like the *SUP2* gene product of *S. cerevisiae*, the Sup2 protein of *P. pinus* represents a fusion of a unique N-terminal part and a region homologous to EF-1α. The comparison of amino acid sequences of the Sup2 proteins reveals high conservation (76%) of the C-terminal region and low conservation (36%) of the N-terminal part where, in addition, the homologous correspondence is ambiguous.

Proteins related to the Sup2 of *S. cerevisiae* were found in *P. pinus* and some other yeast species by the immunoblotting technique.

The relation between the evolutionary conservation of different regions of the Sup2 protein and their functional significance is discussed.

KEY WORDS — Omnipotent suppressor; gene structure; evolutionary conservation; codon bias analysis; *Pichia pinus*; yeast.

INTRODUCTION

Among the genes involved in the control of translational accuracy in yeast *S. cerevisiae*, the *SUP2* gene is one of the most thoroughly studied. Mutations in this gene are able to suppress all three types of nonsense mutations and give rise to a variety of pleiotropic effects, such as temperature sensitivity, paromomycin sensitivity and respiratory deficiency (reviewed by Surguchov *et al.*, 1984). Mutations with similar properties have also been obtained by other authors, as omnipotent suppressors *sup35* (Hawthorne and Leupold, 1974), *supP* (Gerlach, 1975), frameshift suppressors *suf12* (Culbertson *et al.*, 1982) and allosuppressors *sal3* (Cox, 1977). At present all these genes, except for the *SUPP*, are cloned and the *SUP2* and *SUF12* genes are sequenced. A comparison of their restriction maps, nucleotide sequences and complementation properties has shown that these genes are identical (Kushnirov *et al.*, 1988; Wilson and Culbertson, 1988; Tuite *et al.*, 1988). Recently a mutation blocking transition from G1 to S phase of the cell cycle was obtained in the *GST1* gene. The data on its localization and nucleotide sequence have shown its identity to the *SUP2* gene (Kikuchi *et al.*, 1988).

Analysis of the nucleotide sequence of the *SUP2* gene (Kushnirov *et al.*, 1987, 1988; Wilson and Culbertson, 1988) revealed a significant homology to the translation elongation factor EF-1α, suggesting that its product represents a previously unidentified factor of translation. It is supposed from codon bias analysis that the Sup2 protein is much less abundant in the cell than ribosomal proteins or EF-1α. The Sup2 protein is essential for viability

‡Author to whom correspondence should be addressed.

0749–503X/90/060461–12 $06.00

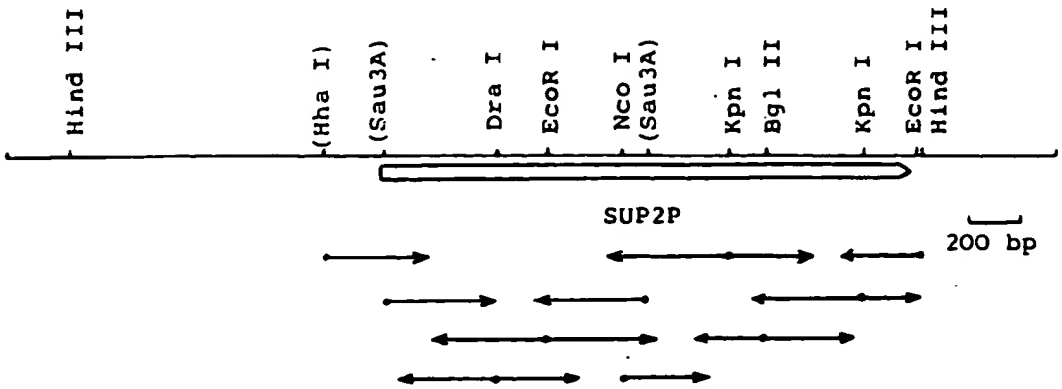


Figure 1. Restriction map and sequencing strategy for the insert of *P. pinus* chromosomal DNA in plasmid pTR30-1 containing the *SUP2* gene analog. The position and orientation of the open reading frame is represented by an arrowed open bar. Small arrows indicate the direction and extent of sequence determination on individual clones. Not all sites for *Sau*3AI and *Hha*I are shown.

and is not interchangeable with EF-1α. Though its function is not established, both involvement in protein synthesis and participation in transition from G1 to S may account for its indispensability to the cell.

We were interested in determining the conservative elements in the *SUP2* gene and evaluating its evolutionary conservation. For this purpose, the functional homologue of the *SUP2* gene from the methylotrophic yeast *Pichia pinus*, taxonomically distant from *S. cerevisiae* (Kreger-van Rij, 1984), was cloned and sequenced. The comparison of deduced amino acid sequences of the Sup2 proteins of *S. cerevisiae* and *P. pinus* reveals strong conservation of the region homologous to EF-1α, especially on the stretch presumably involved in GTP and aminoacyl tRNA binding. The homology level of N-terminal regions, on the contrary, was found to be low.

MATERIALS AND METHODS

Strains and genetic methods

The following strains were used: *S. cerevisiae* 33G-D373 (*MATα, ade2-144,717, pheA10, his7-1, lys9-A21, leu2-3,112, ura3-52, trp1-289*), R183-33G-D373 (*sup2-183 (Ts)* mutant of 33G-D373) and 8H8 (*Matα/Matα, leu2-3,112/leu2-3,112, ura3-52/ura3-52, his3-Δ1/his3-Δ1, trp1-289/trp1-289, SUP2/sup2::URA3*); *E. coli* HB101 and JM103; *P. pinus* MH4 (wild type). *E. coli* cells were transformed as described by Hanahan (1985); transformation of *S. cerevisiae* was performed according to Ito *et al.* (1983). Standard methods of yeast genetics (Sherman *et al.*, 1986) were used. The *SUP2* gene of *S. cerevisiae* in the present work is designated as *SUP2S* and its homologue from *P. pinus* as *SUP2P*.

Construction of genomic library of P. pinus

Chromosomal DNA of *P. pinus* was partially digested by restriction endonuclease *Sau*3A. Fragments of 5 to 15 kb were isolated and inserted in vector YEp13 (Broach *et al.*, 1979), which was cleaved by *Bam*HI and treated with alkaline phosphatase. Ligation mixture was used for transformation of *E. coli* HB101 and approximately 2×10^4 colonies were obtained. Plasmid DNA was isolated from the mixture of these colonies and used for the cloning of the *SUP2P* gene.

RNA analysis

Total yeast RNA was isolated and fractionated on 1·1% agarose gel containing formaldehyde according to Sherman *et al.* (1986), transferred to a Zeta-Probe membrane (Bio-Rad Laboratories) and hybridized under conditions recommended by the membrane manufacturer. RNA blots were probed with single-stranded recombinant M13 phages labeled by extension of hybridization probe primer. Three M13 clones were used containing: (1) *Bcn*I-*Xba*I fragment including the entire *SUP2S* gene with 200 bp 5′ and 3′ flanking regions, (2) *Bcn*I-*Hpa*I fragment including N-terminal A and B regions of the *SUP2S* gene (see Figure 5) and (3) 1·5 kb *Eco*RI fragment of the *SUP2P* gene.

DNA sequencing and analysis of the sequence

Restriction fragments of the *SUP2P* gene were recloned in plasmids pTZ18/19R (Pharmacia). The

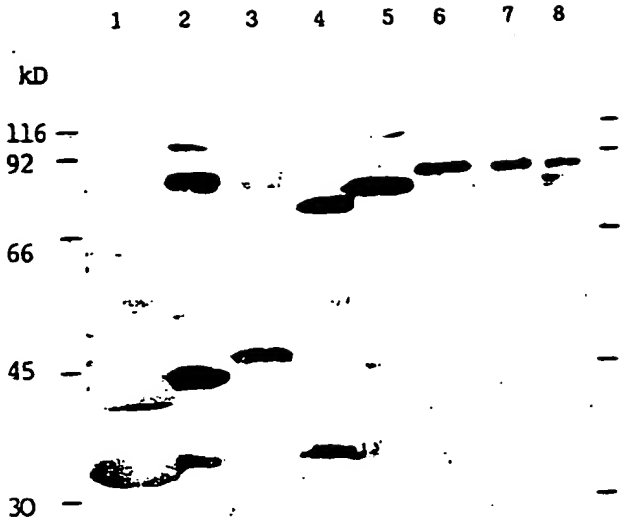


Figure 2. Immunoblot analysis of different yeast species. Using rabbit antisera to the fusion protein of SUP2S and β-galactosidase, cell lysates of the following yeast species were analysed: (1) *Hansenula polymorpha*; (2) *Candida maltosa*; (3) *Candida utilis*; (4) *Kluyveromyces lactis*; (5) *S. cerevisiae*; (6) *P. pinus*; (7) *S. cerevisiae* haploid segregant containing *SUP2S* gene disruption and *SUP2P* gene on pTR30-1 plasmid; (8) same, as (7), but with intact *SUP2S* gene.

single-stranded form of the plasmids was obtained after superinfection with M13K07 phage of *E. coli* strain JM103 carrying these plasmids and used for sequencing by the dideoxy method of Sanger *et al.* (1977). Analysis of the nucleotide and amino acid sequences was accomplished with the use of the Microgenie® software package (Beckman Instruments).

Immunoblot analysis

Rabbit antiserum against fusion protein of Sup2S and β-galactosidase purified from *E. coli* was used (Didichenko *et al.*, submitted). Yeast whole-cell extracts were prepared as described by Last and Woolford (1986). Denatured proteins were separated on 7% polyacrylamide gel containing SDS (Laemmli, 1970) and electrophoretically transferred to nitrocellulose sheets (Towbin *et al.*, 1979). Antigen-antibody reactions and visualization of those complexes with peroxidase-coupled goat anti-rabbit serum were carried out as described by Allis *et al.* (1984).

RESULTS AND DISCUSSION

Cloning of the SUP2 *gene of* P. pinus

S. cerevisiae strain R183-33G-D373, carrying the *sup2* mutation temperature sensitive to 36°C, was transformed with the genomic library of *P. pinus*. Transformants prototrophic for leucine were obtained at 20°C and then transferred to 36°C. Among several thousands of transformants, 30 colonies able to grow at non-permissive temperature (Htr) were selected. Six of them lost Htr phenotype at non-selective medium concurrently with the plasmid marker *LEU2*, indicating that temperature-sensitivity depends on the presence of plasmid in these cases.

Plasmid DNA was isolated from two Htr transformants. Both plasmids carried inserts identical in

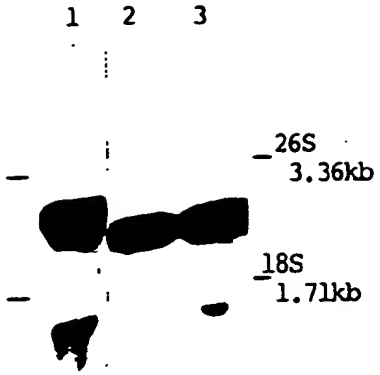


Figure 3. Identification of *SUP2* transcripts. 20 μg of total RNA of *P. pinus* (lane 1) or *S. cerevisiae* (lanes 2 and 3) were analysed by Northern blot hybridization using the M13 probes carrying: 1·5 kb *Eco*RI fragment of the *SUP2P* gene (lane 1), *Bcn*I-*Hpa*I fragment containing the N-terminal part of the *SUP2S* gene non-homologous to EF-1α (lane 2), *Bcn*I-*Xba*I fragment including the entire *SUP2S* gene with 200 bp 5′ and 3′ flanking regions (lane 3). Positions of ribosomal RNAs are indicated as size markers.

size (4·5 kb) and restriction map (Figure 1). One of these plasmids, designated pTR30-1, was used in further work.

Functional analysis of the cloned DNA fragment

To check the ability of the cloned DNA fragment to substitute functionally for the *SUP2S* gene, diploid 8H8 heterozygous for *sup2::URA3* allele (Ter-Avanesyan *et al.*, 1989) was transformed with pTR30-1 plasmid. One of the transformants was subjected to tetrad analysis. Six tetrads, where the plasmid was lost, segregated 2:2 for viability, which is typical for diploid 8H8; besides, there were no viable Ura$^+$ spores (containing *sup2* disrupted allele). In 17 out of 19 tetrads carrying the plasmid, all four spores were viable and segregated 2Ura$^+$: 2Ura$^-$. Two tetrads had each three viable spores, one of them being Ura$^+$. One of the Ura$^+$ haploid clones was analysed using the Western blot technique (Figure 2). One can see that in this clone the Sup2P protein band is present, while the Sup2S band is absent. Thus, plasmid pTR30-1 is able to maintain viability of *S. cerevisiae* cells in the absence of the functional Sup2S product and, therefore, it contains a gene that is functionally equivalent to *SUP2S* gene.

It was shown recently that plasmid-mediated amplification of the wild-type *SUP2S* gene leads to suppression of all three types of nonsense mutations in *S. cerevisiae* (Chernoff *et al.*, 1988). The relation between the mechanisms of this phenomenon and the suppressor effect of the *sup2* mutations is not clear yet. Multicopy plasmid pTR30-1 carrying the wild-type allele of the *SUP2P* gene is also able to cause suppression of nonsense mutations in *S. cerevisiae*, although less efficiently than plasmids with the *SUP2S* gene. In particular, pTR30-1 causes suppression of ochre mutations *his7-1* and *lys9-A21* in the strain 33G-D373. Growth of transformants of this strain with plasmid carrying the *SUP2S* gene, on media without histidine at 30°C, was seen on the 6th day, while transformants with plasmid carrying the *SUP2P* gene grew on this media on the 11th day.

Transcripts of the SUP2 *genes*

Two transcripts were found for the *SUP2* gene in both *S. cerevisiae* and *P. pinus*, with the larger transcripts being five to ten times more abundant than the smaller ones (Figure 3). The larger transcript of *SUP2S* is 2·3 kb long and includes a complete coding sequence of the gene (Kushnirov *et al.*, 1988). Probably, the same is true for the 2·6 kb transcript of *SUP2P*. Smaller transcripts of both *SUP2S* and *SUP2P* have a size of 1·4 kb. These mRNAs are not the transcripts coding for EF-1α, which have approximately the same size and 51% homology to the *SUP2*, since they were observed even at high stringency hybridization, when the *SUP2P*/*SUP2S* cross-hybridization (76% homology) was not detected. The 1·4 kb mRNA of *SUP2S* does not hybridize to a probe containing the entire N-terminal part of the gene non-homologous to EF-1α. The size of the region homologous to EF-1α is around 1·3 kb for both *SUP2S* and *SUP2P* genes. Therefore, we suggest that only this region is included in 1·4 kb RNA in both *SUP2S* and *SUP2P*.

Nucleotide sequence of the SUP2P *gene*

A set of restriction fragments of the *SUP2P* gene from plasmid pTR30-1 was recloned into vector pTZ18/19R and sequenced by the dideoxy method of Sanger *et al.* (1977) (Figure 1). The sequence obtained (Figure 4) contains a single long open reading frame of 2223 bp able to code for a protein

```
                                                                                                                    CCACGTATATTTCTGAAAAATTTTCAGATTTCTTAACTAA
-160
-120  TAGATGAAACTCCAATTATTTTTCGTTACTTGATGAATTTAAGTTGATTCACTTTTATTAGTGTACGTACCCCTTGTTGAACATAAATGGTCTTTTAACTCGCCATCAAGCACCCAAAAG

   1  ATG TCT CAA GAT CAA CAG CAA CAG CAA CAG TTT AAT GCC AAT AAC TTG GCT GGC AAT GTT CAA AAC ATC AAC TTA AAT GCT CCA GCT TAC
   1  Met Ser Gln Asp Gln Gln Gln Gln Gln Phe Asn Ala Asn Asn Leu Ala Gly Asn Val Gln Asn Ile Asn Leu Asn Ala Pro Ala Tyr

  90  GAC CCT GCC GTT CAA TCT TAT ATT CCA AAC ACT GCC CAA GCA TTT GTT CCC TCT GCT CAG CCA TAC ATT CCA GGC CAG CAG GAA CAA CAG
  30  Asp Pro Ala Val Gln Ser Tyr Ile Pro Asn Thr Ala Gln Ala Phe Val Pro Ser Ala Gln Pro Tyr Ile Pro Gly Gln Gln Glu Gln Gln

 180  TTT GGT CAA TAT GGT CAG CAA CAG CAA AAT TAC AAC CAA GGT GGC TAC AAC AAT TAC AAC AAC AGG GGT GGT TAC AGC AAC AAC AGA GGT
  60  Phe Gly Gln Tyr Gly Gln Gln Gln Gln Asn Tyr Asn Gln Gly Gly Tyr Asn Asn Tyr Asn Asn Arg Gly Gly Tyr Ser Asn Asn Arg Gly

 270  GGC TAC AAC AAC AGC AAC AGA GGG GGC TAT AGC AAC TAC AAC AGC TAT AAC ACC AAC AGC AAC CAA GGT GGT TAT AGT AAT TAC AAC AAC
  90  Gly Tyr Asn Asn Ser Asn Arg Gly Gly Tyr Ser Asn Tyr Asn Ser Tyr Asn Thr Asn Ser Asn Gln Gly Gly Tyr Ser Asn Tyr Asn Asn

 360  AAT TAC GCC AAC AAC AGC TAC AAT AAT AAT AAT AAC TAT AAC AAC AAC TAC AAT CAA GGT TAT AAT AAT TAC AAC AGC CAA CCC CAA GGT
 120  Asn Tyr Ala Asn Asn Ser Tyr Asn Asn Asn Asn Asn Tyr Asn Asn Asn Tyr Asn Gln Gly Tyr Asn Asn Tyr Asn Ser Gln Pro Gln Gly

 450  CAA GAC CAA CAA CAA GAG ACC GGT TCC GGT CAA ATG TCT TTA GAG GAC TAC CAA AAA CAG CAA AAG GAA AGT TTG AAC AAA CTG AAC ACC
 150  Gln Asp Gln Gln Gln Glu Thr Gly Ser Gly Gln Met Ser Leu Glu Asp Tyr Gln Lys Gln Gln Lys Glu Ser Leu Asn Lys Leu Asn Thr

 540  AAA CCA AAG AAG GTT TTA AAG TTA AAC TTG AAC TCA AGT ACT GTC AAG GCA CCA ATT GTT ACC AAA AAG AAG GAA GAA GAA CCT GTC AAT
 180  Lys Pro Lys Lys Val Leu Lys Leu Asn Leu Asn Ser Ser Thr Val Lys Ala Pro Ile Val Thr Lys Lys Lys Glu Glu Glu Pro Val Asn

 630  CAA GAA AGT AAG ACC GAA GAA CCG GCT AAA GAA GAA ATC AAG AAC CAA GAG CCA GCT GAA GCA GAA AAT AAG GTT GAA GAA GAG TCA AAG
 210  Gln Glu Ser Lys Thr Glu Glu Pro Ala Lys Glu Glu Ile Lys Asn Gln Glu Pro Ala Glu Ala Glu Asn Lys Val Glu Glu Glu Ser Lys

 720  GTT GAA GCC CCA ACT GCT GCT AAG CCA GTC AGT GAA TCC GAA TTC CCA GCT TCA ACT CCA AAA ACT GAA GCC AAG GCA AGT AAA GAA GTT
 240  Val Glu Ala Pro Thr Ala Ala Lys Pro Val Ser Glu Ser Glu Phe Pro Ala Ser Thr Pro Lys Thr Glu Ala Lys Ala Ser Lys Glu Val

 810  GCA GCT GCC GCT GCT GCT CTC AAG AAG GAA GTT TCT CAA GCT AAG AAG GAA AGT AAC GTT ACC AAC GCT GAT GCC TTA GTC AAG GAG CAA
 270  Ala Ala Ala Ala Ala Ala Leu Lys Lys Glu Val Ser Gln Ala Lys Lys Glu Ser Asn Val Thr Asn Ala Asp Ala Leu Val Lys Glu Gln

 900  GAG GAG CAA ATT GAT GCC TCC ATT GTC AAC GAT ATG TTC GGT GGT AAG GAC CAC ATG TCC ATC ATT TTC ATG GGT CAC GTT GAT GCT GGT
 300  Glu Glu Gln Ile Asp Ala Ser Ile Val Asn Asp Met Phe Gly Gly Lys Asp His Met Ser Ile Ile Phe Met Gly His Val Asp Ala Gly

 990  AAG TCA ACC ATG GGT GGT AAT TTA TTG TTC TTA ACT GGT GCT GTT GAT AAG CGT ACT GTT GAA AAG TAT GAA AGG GAA GCT AAG GAT GCT
 330  Lys Ser Thr Met Gly Gly Asn Leu Leu Phe Leu Thr Gly Ala Val Asp Lys Arg Thr Val Glu Lys Tyr Glu Arg Glu Ala Lys Asp Ala

1080  GGT AGA CAA GGT TGG TAC TTA TCC TGG ATC ATG GAT ACA AAC AAG GAA GAA AGA AAC GAC GGT AAG ACC ATT GAA GTC GGC AAG TCT TAC
 360  Gly Arg Gln Gly Trp Tyr Leu Ser Trp Ile Met Asp Thr Asn Lys Glu Glu Arg Asn Asp Gly Lys Thr Ile Glu Val Gly Lys Ser Tyr

1170  TTC GAA ACG GAC AAG AGA CGT TAC ACC ATT TTA GAT GCC CCA GGA CAT AAG TTG TAT ATT TCC GAA ATG ATC GGT GGT GCT TCT CAA GCC
 390  Phe Glu Thr Asp Lys Arg Arg Tyr Thr Ile Leu Asp Ala Pro Gly His Lys Leu Tyr Ile Ser Glu Met Ile Gly Gly Ala Ser Gln Ala

1260  GAT GTT GGT GTT TTA GTT ATT TCT TCG AGA AAG GGT GAA TAC GAA GCC GGT TTC GAA AGA GGC GGC CAA TCA AGA GAA CAT GCT ATC TTA
 420  Asp Val Gly Val Leu Val Ile Ser Ser Arg Lys Gly Glu Tyr Glu Ala Gly Phe Glu Arg Gly Gly Gln Ser Arg Glu His Ala Ile Leu

1350  GCT AAA ACT CAA GGT GTT AAC AAG TTG ATT GTT GTG ATA AAC AAG ATG GAT GAC CCA ACT GTT AAC TGG TCC AAG GAG AGA TAC GAA GAA
 450  Ala Lys Thr Gln Gly Val Asn Lys Leu Val Val Val Ile Asn Lys Met Asp Asp Pro Thr Val Asn Trp Ser Lys Glu Arg Tyr Glu Glu

1440  TGT ACT ACC AAA TTA GCC ATG TAC TTA AAG GGT GTT GGG TAC CAA AAA GGT GAT GTC TTG TTT ATG CCT GTC TCT GGA TAT ACT GGC GCT
 480  Cys Thr Thr Lys Leu Ala Met Tyr Leu Lys Gly Val Gly Tyr Gln Lys Gly Asp Val Leu Phe Met Pro Val Ser Gly Tyr Thr Gly Ala

1530  GGT TTG AAA GAA AGG GTC AGT CAA AAA GAT GCT CCA TGG TAC AAC GGC CCA TCA TTA TTA GAA TAC TTA GAC TCC ATG CCA TTG GCC GTT
 510  Gly Leu Lys Glu Arg Val Ser Gln Lys Asp Ala Pro Trp Tyr Asn Gly Pro Ser Leu Leu Glu Tyr Leu Asp Ser Met Pro Leu Ala Val

1620  AGA AAG ATC AAC GAT CCG TTC ATG CTA CCA ATC TCT TCT AAG ATG AAA GAT CTA GGT ACT GTT ATC GAA GGT AAG ATT GAA TCA GGT CAT
 540  Arg Lys Ile Asn Asp Pro Phe Met Leu Pro Ile Ser Ser Lys Met Lys Asp Leu Gly Thr Val Ile Glu Gly Lys Ile Glu Ser Gly His

1710  GTT AAG AAG GGT CAG AAC TTG TTA GTT ATG CCA AAT AAG ACT CAA GTT GAA GTC ACC ACC ATT TAC AAC GAA ACT GAA GCT GAA GCT GAC
 570  Val Lys Lys Gly Gln Asn Leu Leu Val Met Pro Asn Lys Thr Gln Val Glu Val Thr Thr Ile Tyr Asn Glu Thr Glu Ala Glu Ala Asp

1800  AGT GCC TTC TGT GGT GAG CAA GTC AGA CTA AGA CTT AGA GGT ATT GAA GAA GAA GAC CTT TCT GCT GGT TAC GTT TTA TCT TCT ATT AAC
 600  Ser Ala Phe Cys Gly Glu Gln Val Arg Leu Arg Leu Arg Gly Ile Glu Glu Glu Asp Leu Ser Ala Gly Tyr Val Leu Ser Ser Ile Asn

1890  CAC CCA GTT AAG ACA GTT ACC AGA TTT GAA GCC CAA ATC GCC ATT GTT GAA TTA AAG TCT ATT TTA TCT ACT GGT TTC TCA TGT GTT ATG
 631  His Pro Val Lys Thr Val Thr Arg Phe Glu Ala Gln Ile Ala Ile Val Glu Leu Lys Ser Ile Leu Ser Thr Gly Phe Ser Cys Val Met

1981  CAC GTC CAT ACT GCC ATT GAA GAA GTT ACC TTT ACT CAG CTA TTG CAC AAT CTA CAA AAG GGT ACC AAC AGA AGA TCA AAG AAG GCC CCT
 661  His Val His Thr Ala Ile Glu Glu Val Thr Phe Thr Gln Leu Leu His Asn Leu Gln Lys Gly Thr Asn Arg Arg Ser Lys Lys Ala Pro

2071  GCT TTC GCT AAG CAA GGT ATG AAG ATT ATT GCT GTT TTA GAG ACC ACC GAA CCA GTT TGT ATC GAA AGC TAC GAT GAT TAC CCA CAA TTA
 691  Ala Phe Ala Lys Gln Gly Met Lys Ile Ile Ala Val Leu Glu Thr Thr Glu Pro Val Cys Ile Glu Ser Tyr Asp Asp Tyr Pro Gln Leu

2161  GGT AGA TTC ACT TTG AGA GAT CAA GGT CAA ACC ATT GCA ATT GGT AAA GTT ACC AAG CTA TTG TAA ATGTTAATGTCATAATTGTTTGAATTCTGTT
 721  Gly Arg Phe Thr Leu Arg Asp Gln Gly Gln Thr Ile Ala Ile Gly Lys Val Thr Lys Leu Leu ...

2250  TGTCTGTGTGTACGATTACAAGCTT
```

Figure 4. Nucleotide and deduced amino acid sequence of the *SUP2* gene of *P. pinus*. Putative TATA boxes are underlined.

Table 1. Codon usage in *SUP2* genes

		P	S			P	S			P	S			P	S
Phe	TTT	6	9	Ser	TCT	16	12	Tyr	TAT	10	10	Cys	TGT	4	4
Phe	TTC	11	7	Ser	TCC	8	7	Tyr	TAC	27	25	Cys	TGC	0	1
Leu	TTA	22	7	Ser	TCA	9	6		TAA	1	1		TGA	0	0
Leu	TTG	13	15	Ser	TCG	1	3		TAG	0	0	Trp	TGG	4	4
Leu	CTT	2	3	Pro	CCT	4	10	His	CAT	4	7	Arg	CGT	2	6
Leu	CTC	1	0	Pro	CCC	2	2	His	CAC	5	6	Arg	CGC	0	0
Leu	CTA	6	7	Pro	CCA	21	18	Gln	CAA	42	40	Arg	CGA	0	0
Leu	CTG	1	3	Pro	CCG	2	0	Gln	CAG	12	13	Arg	CGG	0	0
Ile	ATT	21	17	Thr	ACT	18	14	Asn	AAT	20	24	Ser	AGT	9	5
Ile	ATC	11	12	Thr	ACC	19	16	Asn	AAC	46	21	Ser	AGC	8	2
Ile	ATA	1	3	Thr	ACA	2	8	Lys	AAA	14	28	Arg	AGA	18	11
Met	ATG	17	19	Thr	ACG	1	1	Lys	AAG	47	38	Arg	AGG	3	1
Val	GTT	35	26	Ala	GCT	30	20	Asp	GAT	18	21	Gly	GGT	44	45
Val	GTC	12	10	Ala	GCC	19	16	Asp	GAC	10	9	Gly	GGC	10	10
Val	GTA	0	9	Ala	GCA	6	7	Glu	GAA	52	44	Gly	GGA	2	3
Val	GTG	1	5	Ala	GCG	0	0	Glu	GAG	10	13	Gly	GGG	2	2

The *SUP2* genes of *P. pinus* and *S. cerevisiae* are designated as 'P' and 'S', respectively. Codons, preferred in highly expressed *S. cerevisiae* genes, are underlined.

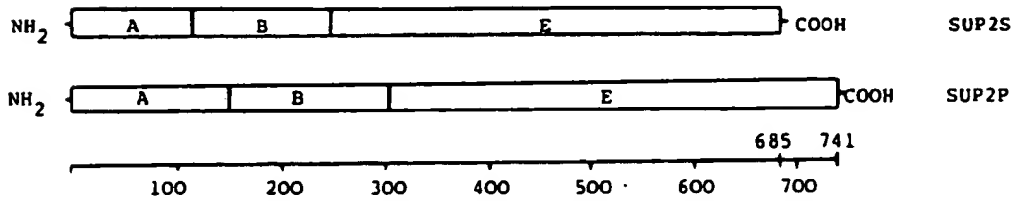


Figure 5. Schematic representation of the predicted primary structures of the *SUP2* proteins of *P. pinus* and *S. cerevisiae*. Each of the proteins may be divided into three regions: A, B and E, beginning from the first, second and third methionine in the sequence, respectively. Region E is homologous to the entire translation elongation factor EF-1α. Regions A and B are distinguished by unusually limited and different amino acid compositions (Table 2). The scale below the scheme is in amino acids.

of 82 367 Da. This exceeds the mass of the Sup2S protein by 5822 Da.

The 5'-flanking region contains three TATAA-like elements, located at position -155(TATATT), -65 (TATTA) and -39 (CATAAA) with respect to the initiating ATG. In the *SUP2S* gene, this element is located in position -104 (TATATT).

Immunoblot analysis of Sup2 proteins

Cell lysates of different yeast species were studied by the immunoblotting procedure with polyclonal antiserum to fusion protein Sup2S-β-galactosidase (Figure 2). An immunoreactive band of 79 kDa was observed in *S. cerevisiae*, and a band of 84 kDa in *P. pinus*. These values correspond well to calculated masses of the Sup2S and Sup2P proteins: 76·5 and 82·4 kDa, respectively. Proteins that react with anti-Sup2S antibody have also been observed in other yeast species. In contrast to *S. cerevisiae* and *P. pinus*, more than one major band has been observed in *Candida utilis*, *Candida maltosa* and *Kluyveromyces lactis*. The reason for the existence of the multiple bands has not been established. However, it is possible to suggest that most of them are related to

Table 2. Amino acid composition of the predicted Sup2 proteins

Amino acids	Region Sup2P A 1–161	B 162–311	E 312–741	The entire Sup2P protein 1–741	Region Sup2S A 1–123	B 124–253	E 253–685	The entire Sup2S protein 1–685
Ala	9	20	26	55	6	9	28	43
Arg	3	0	20	23	2	0	16	18
Asn	42	10	14	66	20	7	18	45
Asp	3	4	21	28	2	7	21	30
Cys	0	0	4	4	0	0	5	5
Gln	30	8	16	54	35	6	12	53
Glu	2	26	34	62	0	23	34	57
Gly	18	0	40	58	21	2	37	60
His	0	0	9	9	0	1	12	13
Ile	3	4	26	33	0	3	29	32
Leu	2	8	35	45	1	7	27	35
Lys	0	25	36	61	1	24	41	66
Met	1	1	15	17	1	1	17	19
Phe	3	1	13	17	3	1	12	16
Pro	7	9	13	29	6	8	16	30
Ser	12	13	26	51	5	10	20	35
Thr	3	8	29	40	0	11	28	39
Trp	0	0	4	4	0	0	4	4
Tyr	20	1	16	37	20	0	15	35
Val	3	12	33	48	0	10	40	50
Total	161	150	430	741	123	130	432	685

Amino acid composition of regions A, B, E and the entire protein is shown for the predicted *Sup2* gene products of *P. pinus* and *S. cerevisiae*. The unusually high content of some amino acids is underlined.

the *SUP2* gene, since the specificity of the antibody is sufficiently high and unrelated proteins are not stained in *S. cerevisiae* and *P. pinus*. The additional bands in *C. utilis* and *C. maltosa* are located around 50 kDa and thus may correspond to short *SUP2* transcripts, similar to 1·4 Kb transcripts found in *S. cerevisiae* and *P. pinus* (see above and Figure 3). Bands at the level of the Sup2S protein are observed in all analysed yeasts, except for *Hansenula polymorpha*. We suggest that these bands represent full-sized Sup2 proteins of particular species.

Codon usage

Codon usage in the *SUP2P* gene is similar to that of the *SUP2S* gene (Table 1). One can see that the same codons are preferably used in this gene as in highly expressed *S. cerevisiae* genes (Bennetzen and Hall, 1982). A single exception may be Leu for which codon TTA is more frequent in *SUP2P*, whereas in *S. cerevisiae* TTG is preferred. The codon adaptation index of the *SUP2P* gene calculated with reference to the highly expressed *S. cerevisiae* genes, according to Sharp and Li (1987), is 0·43 in comparison with 0·33 for the *SUP2S* gene. Codon bias is not uniform along the length of the *SUP2* genes. For regions A, B and E of *SUP2S* (see below for the structure of the *SUP2* genes), the codon adaptation index equals 0·22, 0·30 and 0·37, respectively, and for regions A, B and E of *SUP2P*, it is 0·28, 0·39 and 0·51. A possible explanation for the higher bias level of regions E is their origin from a highly expressed gene, coding for EF-1α.

Codon usage in the *SUP2P* gene is quite different from that observed in another methylotrophic yeast, *Hansenula polymorpha* (Janowicz *et al.*, 1985; Ledeboer *et al.*, 1985), though both species are believed to be closely related (Kurtzman, 1984). *H. polymorpha* also looks different from *P. pinus* and *S. cerevisiae*, when analysed with anti-Sup2S antibody (Figure 2).

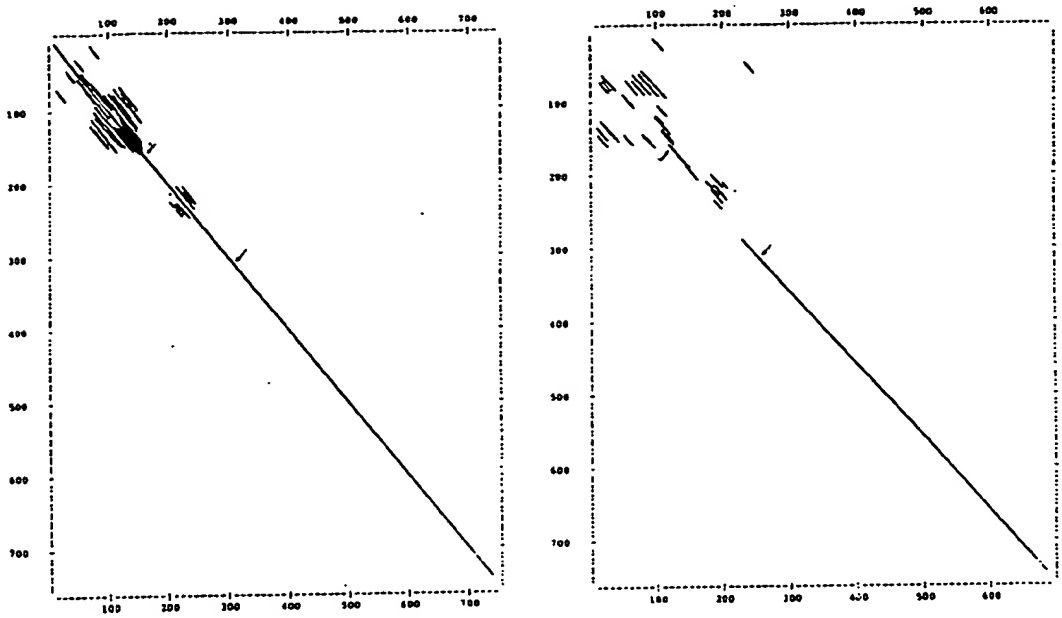


Figure 6. Dot-matrix comparison of Sup2 proteins of *S. cerevisiae* and *P. pinus*: (A) Sup2P with itself; (B) Sup2S and Sup2P. Each dot represents eight exact matches in a window of 20 amino acids. The beginnings of regions B and E (see Figure 5) are marked by arrows.

Structure of the Sup2 proteins

The analysis of the Sup2S sequence revealed that the first three methionines divide it into three regions differing sharply in amino acid composition and predicted secondary structure. The third methionine begins a region homologous to EF-1α. Remarkably, the second and third methionines are conserved in the Sup2P protein and divide it into three regions, similar to those of the Sup2S in their properties (Figure 5, Table 2).

Region A includes amino acids 1–161 in Sup2P or 1–123 in Sup2S and is characterized by an abundance of Asn, Gln, Tyr and Gly, together making 68% in Sup2P and 78% in Sup2S. Secondary structure analysis according to Garnier *et al.* (1978) predicts the existence of a beta structure, but not of an alpha helix.

Region B includes amino acids 162–311 in Sup2P or 124–253 in Sup2S and is rich in charged amino acids, Lys and Glu, together constituting around 35% in each protein. Secondary structure analysis predicts long alpha helical stretches and the absence of beta structure.

Region E includes amino acids 312–741 in Sup2P or 254–685 in Sup2S and is homologous to EF-1α. All amino acids are represented here without any significant bias and all types of secondary structure are encountered.

A characteristic feature of regions A and B, distinguishing them from region E, is the presence of short repeating sequences with high or low levels of homology. Their location in the Sup2P protein may be revealed by dot matrix analysis (Figure 6A). It is obvious that these repeats make up a larger part of region A and a significant part of region B.

A cDNA sequence of the human homologue of the *SUP2* genes, *GST1-Hs*, has been reported recently (Hoshino *et al.*, 1989). The C-terminal part of the predicted Gst1-Hs protein is homologous to EF-1α, while the N-terminal part of 67 amino acids is unique. The N-terminal part of Gst1-Hs is rich in glutamic acid, thus resembling region B of the Sup2 proteins.

Homology of the Sup2 proteins

The alignment of amino acid sequences of the proteins Sup2P, Sup2S (Kushnirov *et al.*, 1988), human Gst1-Hs (Hoshino *et al.*, 1989) and EF-1α of *S. cerevisiae* (Nagashima *et al.*, 1986) is presented

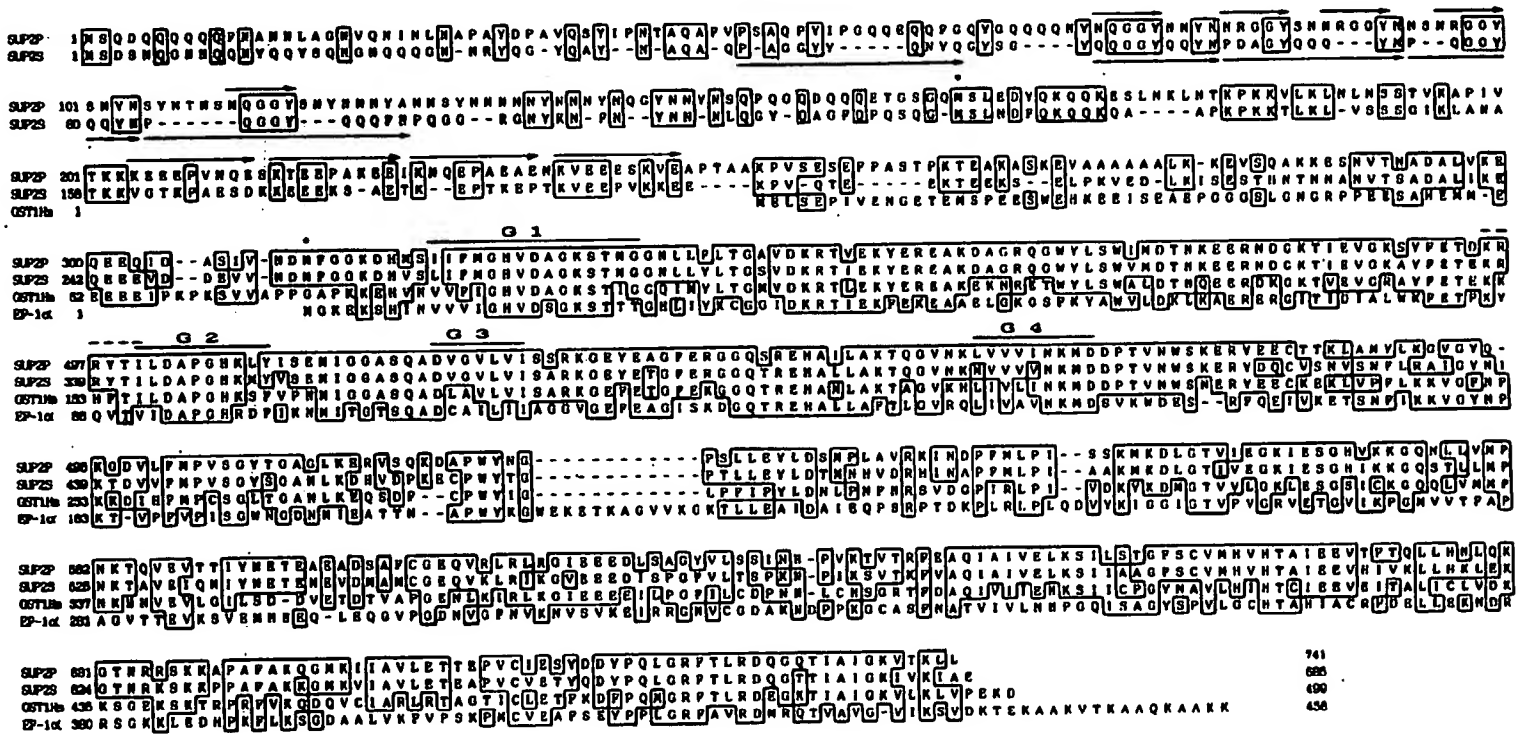


Figure 7. A comparison of amino acid sequences of Sup2 proteins of *P. pinus* and *S. cerevisiae* (Kushnirov *et al.*, 1988), human Gst1-Hs (Hoshino *et al.*, 1989) and EF-1α of *S. cerevisiae* (Nagashima *et al.*, 1986). Sequences were aligned by introducing several gaps (–). Identical amino acids are boxed. Repeating fragments in each sequence are indicated by arrows. Conservative methionines, initiating regions B and E (Figure 5), are marked by asterisks (*). The regions of homology to GTP-binding proteins (G1–G4) are indicated by solid lines, whereas the putative target for cAMP-dependent phosphorylation, sequence KRRYTI (Kikuchi *et al.*, 1988), is shown by a dashed line.

in Figure 7. One can see that the C-terminal part homologous to EF-1α is conserved in the first three proteins. The yeast Sup2 proteins are 76% homologous in this region, whereas their homology to the human Gst1-Hs is around 57%. These values significantly exceed homology of each of these proteins to EF-1α (around 36%). This affirms that the yeast Sup2 proteins and human Gst1-Hs belong to a single protein family, distinct from the EF-1α family. The conservation of the Sup2 proteins is somewhat lower than that of EF-1α. For comparison, homology of EF-1α of *S. cerevisiae* with EF-1α of fungus *Mucor racemosus* (Linz *et al.*, 1986) amounts to 85·5% and with EF-1α of man (Brands *et al.*, 1986) to 81·7%.

The highest conservation is observed in the region of amino acids 312–478 in Sup2P and 254–420 in Sup2S: 90% of exact matches or 100% with conservative substitutions (Dayhoff, 1978). This region corresponds to highly conservative GTP and aminoacyl-tRNA binding domains of EF-1α. A six amino acid sequence, identified by Kikuchi *et al.* (1988) as a potential target site of cAMP-dependent protein kinase (amino acids 337–342 in Sup2S, Figure 7) is identical in the Sup2S and Sup2P proteins. The corresponding sequence in EF-1α differs from Sup2 in four amino acids out of six.

The N-terminal regions A and B are distinguished by the lowest conservation—36%—and the length of these regions differs significantly in the two Sup2 proteins. Moreover, the homologous correspondence is ambiguous here. To establish the cause of the ambiguity, two random amino acid sequences were generated from the N-terminal Sup2S and Sup2P sequences by randomly moving each amino acid within ten residues from its original position. The new sequences have shown 29% homology upon alignment. Thus, the matches of N-terminal sequences of the Sup2 proteins are mainly due to conservation of their very limited amino acid composition and this is the reason for ambiguous conformity of the sequences. The per cent of matches along regions A and B is close to that of the randomized sequences, with the exception of the three fragments: the repeating structure in region A, the beginning and the end of region B (amino acids 55–88, 124–160 and 231–253 in the Sup2S sequence). These fragments are comparatively more conservative (44%, 52% and 70%, respectively) and match each other unambiguously in the Sup2S and Sup2P sequences.

A dot matrix comparison of the Sup2 proteins illustrates the character of similarity of their different regions (Figure 6B). In the C-terminal part the homology is represented by a single straight line. In the N-terminal part the homologous stretches are short and non-colinear.

A comparison of the yeast Sup2 proteins with human Gst1-Hs also demonstrates high variability of their N-terminal part. This part of the Gst1-Hs is much shorter—only 67 amino acids long—and is not homologous to the Sup2 proteins (with the exception of the stretch of 20 residues, adjacent to the EF-1α-like region, Figure 7).

Non-conservation of repeating fragments

Regions A and B of the Sup2S protein include a number of short repeats, the most significant of which is a sequence Gln-Gly-Gly-Tyr-Gln-(Gln)-Gln-Tyr-Asn-Pro repeated about four times in succession (amino acids 57–101, Figure 7) with a high level of conservation (93%). Corresponding to these repeats in the Sup2P protein are irregularly repeated pentapeptides Asn-Gln(Arg)-Gly-Gly-Tyr. In region B of Sup2P there is a contiguous repeat of ten amino acids with low levels of conservation (amino acids 204–243, Figure 7). The corresponding region of Sup2S is not organized in the repeating structure. Thus, repeats are not an indispensable and evolutionary conserved feature of the Sup2 sequences, but probably represent only preferable forms of structural organization.

Functional role and conservation of different regions of SUP2 *genes*

The comparison of the two Sup2 proteins revealed a high conservation level of region E (homologous to EF-1α), whereas the homology of N-terminal parts was found to be low.

Recently we have established by deletion analysis, that the N-terminal part of the Sup2S protein up to Met-254 is not essential for viability (Kushnirov *et al.*, in press). At the same time, the region homologous to EF-1α is essential and sufficient for viability. Thus, one can find a correlation between functional importance and conservation of the structure: the vitally important part of the Sup2 protein is highly conserved, while the region that is non-essential for viability possesses low conservation.

REFERENCES

Allis, C. D., Allen, R. L., Wiggens, J. C., Chicoine, L. G. and Richman, R. (1984). Proteolytic processing of H1-like histones in chromatin: a physiologically and

developmentally regulated event in *Tetrahymena* micronuclei. *J. Cell Biol.* **99**, 1669–1677.

Bennetzen, J. L. and Hall, B. D. (1982). Codon selection in yeast. *J. Biol. Chem.* **257**, 3026–3031.

Brands, J. H. G. M., Maassen, J. A., van Hemert, F. J., Amons, R. and Möller, W. (1986). The primary structure of the α subunit of human elongation factor 1; structural aspects of guanine-nucleotide-binding sites. *Eur. J. Biochem.* **155**, 167–171.

Broach, J. R., Strathern, J. N. and Hicks, J. B. (1979). Transformation in yeast—development of a hybrid cloning vector and isolation of the *CAN1* gene. *Gene* **8**, 121–133.

Chernoff, Yu. O., Derkach, I. L., Dagkesamanskaya, A. R., Tikhomirova, V. L., Ter-Avanesyan, M. D. and Inge-Vechtomov, S. G. (1988). Nonsense-suppression upon amplification of the gene, coding for a protein factor of translation. *Doklady Akad. Nauk USSR* **301**, 1227–1229. (Proceedings of the Academy of Sciences of the USSR, in Russian.)

Cox, B. S. (1977). Allosuppressors in yeast. *Genet. Res.* **30**, 187–205.

Culbertson, M. R., Gaber, R. F. and Cummins, C. M. (1982). Frameshift suppression in *Saccharomyces cerevisiae*. V. Isolation and genetic properties of nongroup-specific suppressors. *Genetics* **102**, 361–378.

Dayhoff, M. O. (Ed.) (1978). *Atlas of Protein Sequence and Structure*, vol. 5, suppl. 1 to 3. National Biomedical Research Foundation, Washington, DC.

Garnier, J., Osguthorpe, D. J. and Robson, B. (1978). Analysis of accuracy and implications of simple methods for predicting secondary structure of globular proteins. *J. Mol. Biol.* **120**, 97–120.

Gerlach, W. L. (1975). Mutational properties of some amber-ochre super-suppressors in *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* **138**, 53–63.

Hanahan, D. (1985). Techniques for transformation of *E. coli*. In Glover, D. M. (Ed.), *DNA Cloning: a Practical Approuch*, vol. 1. IRL Press, Oxford, Washington, DC, pp. 109–135.

Hawthorne, D. C. and Leupold, U. (1974). Suppressor mutations in yeast. *Curr. Top. Microbiol. Immunol.* **64**, 1–47.

Hoshino, S.-I., Miyazawa, H., Enomoto, T., Hanaoka, F., Kikuchi, Y., Kikuchi, A. and Ui, M. (1989). A human homologue of the yeast *GST1* gene codes for a GTP-binding protein and is expressed in a proliferation-dependent manner in mammalian cells. *EMBO J.* **8**, 3807–3814.

Ito, H., Fukuda, Y., Murata, K. and Kimura, A. (1983). Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* **153**, 163–168.

Janowicz, Z. A., Eckart, M. R., Drewke, C., Roggenkamp, R. O. and Hollenberg, C. P. (1985). Cloning and characterization of the *DAS* gene, encoding the major methanol assimilatory enzyme from the methylotrophic yeast *Hansenula polymorpha*. *Nucl. Acids Res.* **13**, 3043–3062.

Kikuchi, Y., Shimatake, H. and Kikuchi, A. (1988). A yeast gene required for the G1 to S transition encodes a protein containing an A-kinase target site and GTPase domain. *EMBO J.* **7**, 1175–1182.

Kreger-van Rij, N. J. V. (Ed.) (1984). *The Yeasts—Taxonomic Study*, 3rd edn. Elsevier Science Publishers, Amsterdam.

Kurtzman, C. P. (1984). Synonymy of the yeast genera *Hansenula* and *Pichia* demonstrated through comparisons of deoxyribonucleic acids relatedness. *Antonie van Leeuwenhoek* **50**, 209–217.

Kushnirov, V. V., Ter-Avanesyan, M. D., Surguchov, A. P., Smirnov, V. N. and Inge-Vechtomov, S. G. (1987). Localization of possible functional domains in *SUP2* gene product of the yeast *Saccharomyces cerevisiae*. *FEBS Lett.* **215**, 257–260.

Kushnirov, V. V., Ter-Avanesyan, M. D., Telckov, M. V., Surguchov, A. P., Smirnov, V. N. and Inge-Vechtomov, S. G. (1988). Nucleotide sequence of the *SUP2(SUP35)* gene of *Saccharomyces cerevisiae*. *Gene* **66**, 45–54.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

Last, R. L. and Woolford, J. L. (1986). Identification and nuclear localization of yeast pre-messenger RNA processing components: RNA2 and RNA3 proteins. *J. Cell Biol.* **103**, 2103–2112.

Ledeboer, A. M., Edens, L., Maat, J., Visser, C., Bos, J. W., Verrips, C. T., Janowicz, Z., Eckart, M., Roggenkamp, R. and Hollenberg, C. P. (1985). Molecular cloning and characterization of a gene, coding for methanol oxidase in *Hansenula polymorpha*. *Nucl. Acids Res.* **13**, 3063–3082.

Linz, J. E., Lira, L. M. and Sypherd, P. S. (1986). The primary structure and the functional domains of an elongation factor 1 from *Mucor racemosus*. *J. Biol. Chem.* **261**, 15022–15029.

Nagashima, K., Kasai, M., Nagata, S. and Kaziro, Y. (1986). Structure of the two genes coding for polypeptide chain elongation factor 1 from *Saccharomyces cerevisiae*. *Gene* **45**, 265–273.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467.

Sharp, P. M. and Li, W.-H. (1987). The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucl. Acids Res.* **15**, 1281–1293.

Sherman, F., Fink, G. R. and Hicks, J. B. (1986). *Methods in Yeast Genetics*. Cold Spring Harbor Laboratory, New York.

Surguchov, A. P., Smirnov, V. N., Ter-Avanesyan, M. D. and Inge-Vechtomov, S. G. (1984). Ribosomal suppression in eukaryotes. *Physicochem. Biol. Rev.* **4**, 147–205.

Ter-Avanesyan, M. D., Dagkesamanskaya, A. R. and Smirnov, V. N. (1989). Null alleles of the *SUP1* and *SUP2* genes: interaction with cytoplasmic

determinants. *Doklady Akad. Nauk USSR* 308, 1472–1475. (Proceedings of Academy of Sciences of the USSR, in Russian.)

Towbin, H., Staehelin, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* 76, 4350–4354.

Tuite, M. F., Izgu, F., Grant, C. M. and Crouzet, M. (1988). Allosuppressor genes and the control of translational fidelity in yeast. *Biochem. Soc. Transact.* 16, 1988–1989.

Wilson, P. G. and Culbertson, M. R. (1988). SUF12 suppressor protein of yeast: a fusion protein related to the EF-1α family of elongation factors. *J. Mol. Biol.* 199, 559–573.

J. Mol. Biol. (1988) 199, 559–573

SUF12 Suppressor Protein of Yeast

A Fusion Protein Related to the EF-1 Family of Elongation Factors

Patricia G. Wilson and Michael R. Culbertson†

Laboratories of Genetics and Molecular Biology
University of Wisconsin
Madison, WI 53706, U.S.A.

(Received 17 April 1987, and in revised form 24 August 1987)

Mutations at the *suf12* locus were isolated in *Saccharomyces cerevisiae* as extragenic suppressors of +1 frameshift mutations in glycine (GGX) and proline (CCX) codons, as well as UGA and UAG nonsense mutations. To identify the SUF12 function in translation and to understand the relationship between *suf12*-mediated misreading and translational frameshifting, we have isolated an $SUF12^+$ clone from a centromeric plasmid library by complementation. $SUF12^+$ is an essential, single-copy gene that is identical with the omnipotent suppressor gene $SUP35^+$. The $2{\cdot}3 \times 10^3$ base $SUF12^+$ transcript contains an open reading frame sufficient to encode a 88×10^3 M_r protein. The pattern of codon usage and transcript abundance suggests that $SUF12^+$ is not a highly expressed gene.

The linear SUF12 amino acid sequence suggests that SUF12 has evolved as a fusion protein of unique N-terminal domains fused to domains that exhibit essentially co-linear homology to the EF-1 family of elongation factors. Beginning internally at amino acid 254, homology is more extensive between the SUF12 protein and EF-1α of yeast (36% identity; 65% with conservative substitutions) than between EF-1α of yeast and EF-Tu of *Escherichia coli*. The most extensive regions of SUF12/EF-1α homology are those regions that have been conserved in the EF-1 family, including domains involved in GTP and tRNA binding.

It is clear that SUF12 and EF-1α are not functionally equivalent, since both are essential *in vivo*. The N-terminal domains of SUF12 are unique and may reflect, in part, the functional distinction between these proteins. These domains exhibit unusual amino acid composition and extensive repeated structure.

The behavior of *suf12-null*/$SUF12^+$ heterozygotes indicates that *suf12* is co-dominantly expressed and suggests that *suf12* allele-specific suppression may result from functionally distinct mutant proteins rather than variation in residual wild-type $SUF12^+$ activity. We propose a model of *suf12*-mediated frameshift and nonsense suppression that is based on a primary defect in the normal process of codon recognition.

1. Introduction

Genetic analyses of translational frameshift suppressors in *Escherichia coli* indicate that defects in the translational machinery that result in translational frameshifting are also associated with errors in codon recognition. Translational suppressors of frameshift and nonsense mutations include the S4 *ram* mutant (ribosomal ambiguity mutations; Atkins *et al.*, 1972), the *supK* tRNA methylase mutant (Atkins & Ryce, 1974), and the *ksgA* mutant, which is defective for a methytransferase specific for 16 S ribosomal RNA (van Buul *et al.*, 1984). Translational suppressors that are limited to frameshift suppression have not been identified. These observations suggest that the translational reading frame is determined by codon–anticodon interactions (Lipmann, 1969; Gaber & Culbertson, 1984; Curran & Yarus, 1986) and maintained by the accuracy of tRNA selection.

In *E. coli* translational accuracy is dependent on ribosome : ternary complex (aminoacyl-tRNA · EF-Tu · GTP) interactions. According to the "Internal Standard Hypothesis" (Thompson *et al.*, 1986), the relative rates of EF-Tu · GDP and aminoacyl-tRNA disassociation from the ribosome are the critical determinants of non-cognate tRNA rejection. Factors that influence these rates may also

† Author for correspondence at: Laboratory of Molecular Biology and Biophysics, 1525 Linden Drive, Madison, WI 53706, U.S.A.

0022-2836/88/040559-15 $03.00/0

influence translational accuracy. Consistent with this model, altered forms of EF-Tu (Vijgenboom *et al.*, 1985; Tapio & Kurland, 1986) as well as the ribosomal proteins required for efficient GTP hydrolysis (Kirsebom & Isaksson, 1985) and ternary complex binding (Changehien *et al.*, 1978) have been identified as nonsense suppressors.

The molecular mechanisms that control translational accuracy in eukaryotes are not as well characterized as those of *E. coli* (Moldave, 1985). However, it is clear that aminoacyl-tRNA is delivered to the eukaryotic ribosome in a ternary complex with EF-1α, the functional analogue of EF-Tu, and GTP. The most extensively studied fidelity mutants in eukaryotes are the "omnipotent" *sup35* and *sup45* suppressors of *Saccharomyces cerevisiae*. These mutations were isolated as recessive, non-specific nonsense suppressors (Hawthorne & Leupold, 1974) and are probably allelic to *sup2* and *sup1* (Surguchov *et al.*, 1984), and *supP* and *supQ* (Gerlach, 1975), respectively. The *SUP45*$^+$ gene encodes an essential 49×10^3 M_r acidic protein that is probably not present in amounts stoichiometric with the ribosome (Himmelfarb *et al.*, 1985; Breining & Piepersberg, 1986). *SUP45*$^+$ transcription appears to be co-ordinately regulated with the transcription of other genes of the translational machinery (Himmelfarb *et al.*, 1985; Breining & Piepersberg, 1986). Although genetic and molecular analyses suggest that SUP45 is a component of the translational machinery, the function of this protein in translation is not clear.

In previous studies of frameshift suppression in the yeast *S. cerevisiae*, *suf12* mutations were isolated as extragenic suppressors of +1 frameshift mutations in either glycine (GGX) or proline (CCX) codons (Culbertson *et al.*, 1982). Further genetic analysis indicated that these suppressors are not limited to frameshift suppression, since *suf12* alleles also suppress UGA and UAG nonsense mutations. Genetic mapping suggested that *suf12* may be allelic to *sup35*. Suppression is recessive but varies with the allele of *suf12* examined, suggesting that the *suf12* defect is not a simple loss of function.

To understand the role of the *SUF12*$^+$ gene product in codon recognition and maintenance of the translational reading frame, we have isolated the *SUF12*$^+$ gene by complementation. The *SUF12*$^+$ gene is an essential, single-copy gene that is allelic to *SUP35*$^+$. Our results suggest that SUF12 is a soluble factor that has evolved as a fusion protein of unique N-terminal domains fused to domains related to EF-1α.

Table 1
Yeast and bacterial strains used

	Genotype	Source
Yeast strains		
S288C	a: *CUP1, gal1, mal1, SUC2*	R. K. Mortimer
PW108-4C	α: *his4-713, leu2-3, met2-1, suf12-3, ura3-52*	This study
PW118-22A	a: *leu2-3, met2-1, trp1-1, lys2-1, hom2, aro1, ura3-52*	This study
PW145	α/a: *leu2-3/leu2-1, leu2-112, met2-1/met2-1, trp1-1/trp1-Δ, lys2-1/+, hom2/+, aro1/+, ura3-52/ura3-52, his4-713/+*	This study
PW145-3C	α: *suf12::LEU2*$^+$, *leu2-3* or *leu2-3, leu2-112, aro1 hom2, ura3-52, met2-1,* [pPWE.RX]	This study
PY39	α: *leu2-3, trp1-Δ, ura3-52*	P. Leeds
SL797-2C	a: *met2-1, leu2-1, aro7-1, trp1-1, lys2-2 ura3-52, His*$^-$, *sup35-4*	S. Liebman
Bacterial strains		
6507	*recA, pyr23::*Tn*5*, kan^r, pro$^-$, leu$^-$, m$_k^-$, r$_k^-$	D. Botstein
JM109	*Δlac-proAB, relA1, recA1, endA1, gyrA96, hsdR17, supE44, thi,* [F′*traD36, proAB, lacI*^{q}, *lacZΔM15*]	J. Messing

2. Materials and Methods

(a) *Materials*

Acrylamide, bisacrylamide, formamide, urea and the 17-base bacteriophage M13 universal primer were obtained from Bethesda Research Laboratories (BRL). Formamide and glyoxal were deionized with AG501-X8D mixed-bed resin from Bio-Rad. Deoxy- and dideoxynucleotides used in DNA sequence analysis were from PL Biochemicals. Other nucleotide triphosphates were purchased from International Biotechnologies, Inc. The remaining chemicals were obtained from Sigma. Restriction endonucleases, *E. coli* DNA polymerase I Klenow fragment, and bacteriophage T4 ligase were obtained from BRL, Promega-Biotec, New England Biolabs, and International Biotechnologies, Inc. [^{32}P]dATP (3000 Ci/mmol) and [^{35}S]dATP (650 to 1200 Ci/mmol) were from Amersham and New England Nuclear. Nitrocellulose and APT paper were obtained from Schleicher and Schuell. XAR-5 film for autoradiography was obtained from Kodak.

(b) *Strains, plasmids and media*

Yeast and bacterial strains used in this study are listed in Table 1. The *CEN4* library was constructed of random 15 to 20 kb† *Sau*3A fragments of yeast wild-type DNA ligated into the centromeric plasmid YCp50 (Rose *et al.*, 1987). The plasmids YIp5 (Botstein *et al.*, 1979), EC402 and YCp50 were obtained from R. Davis, E. Craig and M. Johnston, respectively. The M13 phages mp18 and mp19 were obtained from Amersham. Plasmids pFS-3 and pFS-13 containing *TEF1*$^+$ and *TEF2*$^+$, respectively, were obtained from P. Phillipsen. The centromeric plasmid library and the plasmid RB8 were obtained from M. Rose and D. Botstein, respectively.

Media for growth and selection of yeast were prepared as described by Gaber & Culbertson (1982). Media for the growth and selection of bacteria and M13 phage were prepared as described in the *M13 Cloning and Sequencing Manual* (Amersham). Standard yeast genetic techniques have been described by Sherman *et al.* (1971). Standard linkage values were derived from tetrad data by using the equation: X (in centimorgans) = 50 (tetratype asci + 6 non-parental ditype asci)/total asci (Perkins, 1949). The

† Abbreviations used: kb, 10^3 bases or base-pairs; bp, base-pairs.

tson *et al.*, 1982). Further genetic
ed that these suppressors are not
shift suppression, since *suf12* alleles
IGA and UAG nonsense mutations.
ng suggested that *suf12* may be
. Suppression is recessive but varies
of *suf12* examined, suggesting that
is not a simple loss of function.
id the role of the $SUF12^+$ gene
on recognition and maintenance of
al reading frame, we have isolated
gene by complementation. The
s an essential, single-copy gene that
$UP35^+$. Our results suggest that
luble factor that has evolved as a
of unique N-terminal domains fused
ted to EF-1α.

Materials and Methods

(a) *Materials*

sacrylamide, formamide, urea and the
phage M13 universal primer were
ethesda Research Laboratories (BRL).
glyoxal were deionized with AG501-X8D
from Bio-Rad. Deoxy- and dideoxy-
n DNA sequence analysis were from PL
ther nucleotide triphosphates were
nternational Biotechnologies, Inc. The
als were obtained from Sigma. Restric-
es, *E. coli* DNA polymerase I Klenow
acteriophage T4 ligase were obtained
ega-Biotec, New England Biolabs, and
Biotechnologies, Inc. [^{32}P]dATP
nd [^{35}S]dATP (650 to 1200 Ci/mmol)
ersham and New England Nuclear.
id APT paper were obtained from
huell. XAR-5 film for autoradiography
n Kodak.

Strains, plasmids and media

erial strains used in this study are listed
EN4 library was constructed of random
3A fragments of yeast wild-type DNA
entromeric plasmid YCp50 (Rose *et al.*,
ids YIp5 (Botstein *et al.*, 1979), EC402
obtained from R. Davis, E. Craig and
pectively. The M13 phages mp18 and
ined from Amersham. Plasmids pFS-3
ining $TEF1^+$ and $TEF2^+$, respectively,
from P. Phillipsen. The centromeric
id the plasmid RB8 were obtained from
otstein, respectively.
th and selection of yeast were prepared
aber & Culbertson (1982). Media for the
tion of bacteria and M13 phage were
ibed in the *M13 Cloning and Sequencing*
ım). Standard yeast genetic techniques
ed by Sherman *et al.* (1971). Standard
re derived from tetrad data by using the
centimorgans) = 50 (tetratype asci + 6
pe asci)/total asci (Perkins, 1949). The

s used: kb, 10^3 bases or base-pairs; bp,

order of markers in multipoint crosses was obtained from linkage analysis of the recombinant asci.

(c) *Nucleic acid isolation*

Large-scale preparations of yeast chromosomal DNA were obtained by the method of Olson *et al.* (1979). Plasmid DNA and small quantities of genomic DNA (100 to 200 μg) suitable for hybridization experiments were recovered from yeast mini-lysates as described by Sherman *et al.* (1982). Mini-preparations of plasmid DNA were isolated as described by Holmes & Quigley (1981). Large-scale preparations of M13 replicative form and single-stranded phage DNA were prepared as described by Yamamoto & Alberts (1970). RNA was isolated from yeast as described by Lindquist (1981).

(d) *Bacterial and yeast transformations*

Bacterial tranformations of strain 6507 were carried out by the method of Mandel & Higa (1970). Procedures involving M13 phage were performed as described in the *M13 Cloning and Sequencing Manual* (Amersham). Yeast strains were transformed by a modification of the lithium acetate method (Ito *et al.*, 1983). Yeast cells were harvested in early log growth (10^6 cells per ml), concentrated 10-fold in 0·1 M-lithium acetate, incubated at 30°C (or room temperature if the strains were temperature-sensitive for growth) for 1·5 to 3 h, and then concentrated 10-fold further in 0·1 M-lithium acetate. Plasmid DNA (2 to 10 μg) and sheared carrier DNA (50 μg calf thymus) were added to 50 μl of concentrated cells. After 0·5 to 3 h at 30°C (or room temperature), 0·5 ml of 40% (w/v) polyethylene glycol was added and the cells were incubated for 2 to 3 h more at 30°C. Cells were heat-shocked at 42°C for 15 min, washed 3 times in water and plated on selective media.

(e) *Plasmid and phage constructions*

Methods involving DNA restriction endonuclease digestion, gel fractionation of restriction fragments and ligation have been described by Maniatis *et al.* (1982).

To map genetically the genomic source of the plasmid insert, the plasmid pPWI.CS was constructed by ligation of the 2·8 kb *Cla*I–*Sal*I fragment of pPW12.1 into the *Cla*I–*Sal*I sites of the integrative vector, YIp5. Since the *Cla*I restriction site was located in YCp50, this plasmid contained 2·5 kb of the pPW12.1 plasmid insert.

To localize the $SUF12^+$ gene within pPW12.1, several deletion derivatives were constructed. The plasmids pPW12.HΔ, pPW12.SΔ and pPW12.BsΔ were constructed by *Hin*dIII, *Sal*I and *Bst*EII endonuclease digestion, respectively, and ligation at 1 μg DNA/ml. One of the deletion endpoints of pPW12.SΔ and pPW12.HΔ was located in the vector YCp50, but the deletion endpoints of pPW12.BsΔ were within the plasmid insert.

To facilitate subcloning and DNA sequence analysis of the $SUF12^+$ gene, pPW.XΔ and pPW.RX were constructed. The plasmid pPW.XΔ was obtained by deletion of 2 *Xho*I fragments of pPW12.1 as described above. pRW.RX was constructed by deletion of 2 *Eco*RI fragments of pPW.XΔ. The plasmids pPWE.RX and pPW12.RX contain the complete $SUF12^+$ gene and were derived from pPW.XΔ and pPW.RX, respectively. pPWE.RX was obtained by ligation of the 6·0 kb *Cla*I–*Xho*I fragment of pPW.XΔ into the *Cla*I–*Sal*I sites of EC402, a derivative of YIp5 that contains the yeast 2 μm origin of replication, pPW12.RX was constructed by subcloning the 4·5 kb *Eco*RI/*Xho*I fragment of pPW.RX into the *Eco*RI–*Sal*I sites of YCp50.

The plasmids used for the transplacement experiments were derived from pPW.RX. The plasmid pPW.RX::$URA3^+$ was constructed by replacement of the 0·33 kb *Hin*dIII fragment of pPW.RX with the *Hin*dIII fragment of RB8 containing the $URA3^+$ gene. pPW.RX::$LEU2^+$ was constructed by replacement of the *Hpa*I–*Hpa*I and *Hpa*I–*Sal*I fragments of pPW.RX with the *Sal*I–*Hpa*I fragment of YEp13 (Broach *et al.*, 1979) containing the $LEU2^+$ gene. Prior to transformation of yeast, these plasmids were digested with *Eco*RI/*Xho*I and *Nsi*I/*Hin*dIII, respectively, to direct integration to the homologous genomic regions (Rothstein, 1983).

The M13 clone mpHS.3 used to obtain single-stranded DNA probes for nucleic acid hybridizations was constructed by ligation of the 1·0 kb *Hin*dIII–*Sal*I fragment of pPW.RX into mp19.

(f) *Nucleic acid hybridization*

DNA restriction fragments fractionated on 0·8% agarose gels were transferred to nitrocellulose by the method of Southern (1975) as modified by Wahl *et al.* (1979). Filters were prehybridized in 10 to 20 ml of 5 × SSC (SSC is 0·15 M-NaCl, 0·015 M-sodium citrate, pH 7), 0·5% polyvinylpyrrolidone, 0·5% Ficol and 100 μg of sheared calf thymus DNA/ml at 65°C for 16 to 48 h. Hybridization conditions differed only by the addition of a 0·5 ml sample of radioactively labeled probe. Filter washes varied in stringency as indicated in the Figure legends and text.

RNA was denatured in a solution of 1 M deionized glyoxal (pH 5·5), 1 × TBE (pH 7·3) at 50°C for 1 h. Denatured RNA (10 to 20 μg) was fractionated in 1 to 1·5% agarose, 1 × TBE (pH 7·3) gels. Recirculation of the buffer was not required to maintain this system below pH 8·0, as required by the glyoxal method of RNA denaturation (Maniatis *et al.*, 1982). RNA was transferred to APT paper by electroelution as directed by the manufacturer. Prehybridization and hybridization conditions were as for DNA hybridizations except that solutions contained 50% formamide and hybridizations were at 42°C.

Plasmid probes were radioactively labeled (2×10^7 to 8×10^7 cts/min per μg DNA) with [^{32}P]dATP by nick-translation as described by Maniatis *et al.* (1982). M13-derived probes (10^8 to 10^9 cts/min per μg) were obtained by second strand synthesis as follows: 0·2 μg of single-stranded phage DNA was hybridized to 2·5 ng of the 17-base M13 universal primer in 10 mM-Tris·HCl (pH 7·3), 10 mM-$MgCl_2$ at 60°C for 30 min and cooled to room temperature. The polymerization reaction contained 50 nM-dXTP (dCTP, dTTP, dGTP), 10 μCi of [^{32}P]dATP (650 mCi/mmol), 2 to 4 units of Klenow fragment in a volume of 20 μl, and was incubated at room temperature for 1 h. Hybridization solutions contained either 2×10^7 to 8×10^7 cts/min of nick-translated probe or 1×10^8 cts/min of single-stranded M13 probe per 10 ml of solution.

(g) *DNA sequence analysis*

DNA sequences were determined by the methods of Sanger *et al.* (1977) and Biggen *et al.* (1983). The sequence from −378 to +2019 in Fig. 3 was obtained from sequence analysis of both DNA strands except for the region between +438 and +756. We were unable to

obtain the appropriate subclones in the reverse orientation, utilizing several different cloning strategies. The remaining sequence was obtained from one strand of multiple, independent and overlapping clones. In each case the sequences were within the first 200 bp from the primer/cloning sites and are unambiguous. All of the cloning junction points were sequenced in overlapping clones to ensure a contiguous sequence. The DNA sequence between −710 to −378 was obtained from DNA sequence analysis of a single strand (data not shown).

Computer programs for DNA and amino acid sequence analysis were provided by the Genetics Computer Group of the University of Wisconsin (see appropriate Figure legends).

3. Results

(a) *Cloning the* SUF12$^+$ *gene*

We used the recessive suppressor phenotype of *suf12* to isolate a clone of the *SUF12*$^+$ gene by complementation. A yeast strain containing *suf12-3*, the suppressible frameshift mutation *his4-713*, and *ura3-52* (PW108-4C, Table 1), was transformed to uracil prototrophy with DNA prepared from a centromeric plasmid library (Rose *et al.*, 1987). Ura$^+$ transformants were screened for complementation and loss of *suf12*-mediated suppression of *his4-713* by replica plating to selective media lacking uracil and histidine, and media lacking only histidine. We identified three transformants with the appropriate phenotype in a screen of approximately 6000 transformants. These three transformants failed to grow under conditions requiring both suppression and maintenance of the plasmid (histidine and uracil prototrophy), but grew under conditions selecting for plasmid loss and restored suppression (histidine prototrophy). These candidates contained identical plasmids, designated pPW12.1, as shown by restriction map analysis (Fig. 1). The smallest subclone retaining the wild-type *SUF12*$^+$ function, pPW12.RX, contained the 4·5 kb *Eco*RI–*Xho*I fragment of pPW12.1 inserted into YCp50 (Fig. 1).

The genomic source of the plasmid insert was genetically mapped to the *suf12* region of chromosome IV (Culbertson *et al.*, 1982). A plasmid, pPWI.CS, containing the 2·8 kb *Cla*I–*Sal*I fragment of pPW12.1 was integrated into the genome of PW108-4C (Table 1) by homologous recombination (Hinnen *et al.*, 1978). An integrant was mated to a strain containing genetic markers tightly linked to *suf12* on chromosome IV (PW118-22A, Table 1). Sporulation and tetrad analysis of a purified diploid indicated that the plasmid integration site, inferred from the Ura$^+$ phenotype conferred by the integrated plasmid, was near the map position previously reported for *suf12* (Table 2; Culbertson *et al.*, 1982). Genetic linkage of the plasmid insert to the region of the *suf12* locus is consistent with a plasmid copy of the *SUF12*$^+$ gene.

(b) *Identification of the* SUF12$^+$ *transcript*

To localize the *SUF12*$^+$ gene within the *Eco*RI–*Xho*I fragment of pPW12.RX, deletion derivatives of pPW12.1 were constructed (pPW12.HΔ,

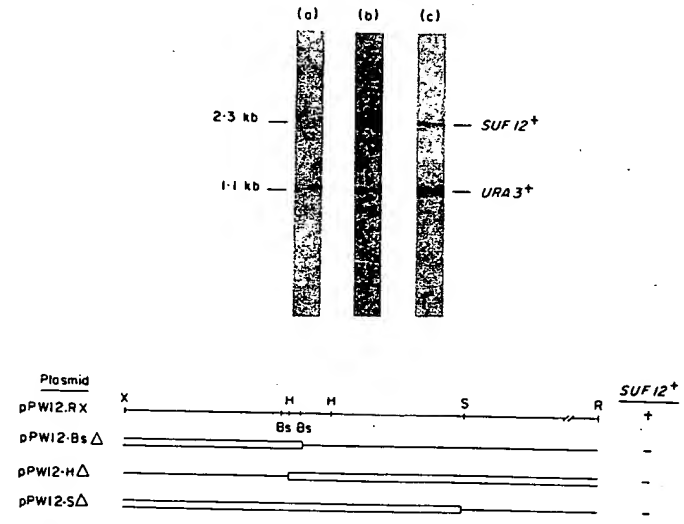


Figure 1. Deletion analysis of pPW12.1: the subclone pPW12.RX contains the 4·5 kb *Eco*RI–*Xho*I fragment of pPW12.1. Deletions of pPW12.1 are represented as open boxes. Complementation analysis of *suf12-3* in PW108-4C (Table 1) indicated that pPW12.RX retains *SUF12*$^+$ function. Since the deletions in pPW12.HΔ and pPW12.BsΔ overlap and each plasmid fails to complement, the region of overlap is probably required for *SUF12*$^+$ expression. Identification of the *SUF12*$^+$ transcript: total wild-type RNA was probed with radioactively labeled (lane a) YCp50 vector (lane b) pPW12.1 (lane c) pPW12.SΔ. The rRNAs were used as standards (Warner, 1982). We could detect hybridization of the plasmid insert to a single 2·3 kb transcript in lanes b and c with stringent filter washes (50% formamide, 55°C, 0·1 × SSC). Bs, *BstE*II; R, *Eco*RI; H, *Hin*dIII; S, *Sal*I; X, *Xho*I.

pPW12.BsΔ and pPW12.SΔ) and screened for their ability to complement *suf12-3* (Fig. 1). A portion of the *Eco*RI–*Xho*I fragment required for *SUF12*$^+$ expression was localized to the 100 bp overlap between the pPW12.HΔ and pPW12.BsΔ deletions, since both plasmids failed to complement *suf12-3* (Fig. 1). Loss of wild-type function by the pPW12.SΔ deletion could be due to loss of all or part of the *SUF12*$^+$ gene.

To correlate the results from deletion mapping with the locations of specific transcripts, total wild-type RNA was probed with radioactively labeled YCp50 (the library vector), pPW12.1 and pPW12.SΔ (Fig. 1(a) to (c)). We detected hybridization of the plasmid insert to a single 2·3 kb transcript, presumably encoded by the *SUF12*$^+$ gene. The intensity of the hybridization signal was approximately equal to that of the *URA3*$^+$ transcript. The plasmid pPW12.SΔ also hybridized to the *SUF12*$^+$ transcript, although the intensity of the hybridization signal was diminished relative to the *URA3*$^+$ transcript. A reduction in the signal relative to the *URA3*$^+$ transcript is consistent with inactivation of the *SUF12*$^+$ gene in pPW12.SΔ by deletion of a region of DNA that normally encodes part of the *SUF12*$^+$ transcript. These results and the deletion analysis of pPW12.1 predict that the pertinent deletion endpoints of pPW12.HΔ, pPW12.BsΔ and pPW12.SΔ are located within the coding and/or regulatory regions of the *SUF12*$^+$ gene. This prediction was confirmed by DNA sequence analysis that revealed a 2055 bp open reading frame containing the respective deletion endpoints (see Fig. 3, below).

(c) SUF12$^+$ *is a single-copy essential gene*

To determine the copy number of the *SUF12*$^+$ gene, wild-type genomic DNA was digested with several enzymes, fractionated by agarose gel electrophoresis, and transferred to nitrocellulose (Southern, 1975). The filter was probed with mpHS.3, a single-stranded M13 clone containing the internal 1·0 kb *Hin*dIII–*Sal*I fragment of the *SUF12*$^+$ coding region (Fig. 2). The presence of a single major band in each gel lane (Fig. 2 lanes a to c) indicates that *SUF12*$^+$ is a single-copy gene.

To determine if *SUF12*$^+$ encodes an essential function, a null allele was constructed *in vitro* by replacing the internal *Hin*dIII fragment of the coding region with the *URA3*$^+$ gene (Fig. 2). The null allele was transplaced (Rothstein, 1983) into PW145, a diploid strain homozygous for *ura3-52*

Table 2
Tetrad analysis

	aro1/hom2	*aro1*/Ura$^+$	*hom2*/Ura$^+$
PD	17	16	19
T	3	5	2
NPD	0	0	0
Total	20	21	21
cM	7·5	11·9	5·2
	aro1/hom2	*aro1/suf12*	*hom2/suf12*
cM	6·5	8·5	3·5

Tetrads were obtained from a mating between PW108-4C[pPWI.CS] and PW118-22A. The plasmid integration site was inferred from the Ura$^+$ phenotype conferred by the *URA3*$^+$ gene of pPWI.CS. The genetic map distances between *URA3*$^+$, *aro1* and *hom2* are similar to those reported for *suf12*, *aro1* and *hom2* (Culbertson *et al.*, 1982). The relative order, *aro1 hom2* Ura$^+$, was determined by analysis of the tetratype (T) asci. This order is similar to the relative map positions of *aro1*, *hom2* and *suf12* (Culbertson *et al.*, 1982). PD, parental ditype; NPD, non-parental type asci; cM, centimorgan.

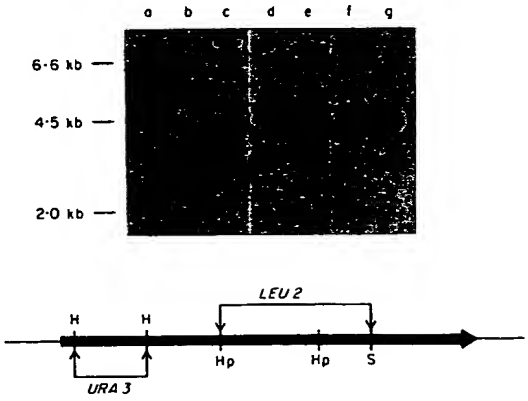


Figure 2. Southern hybridization analysis: genomic yeast DNA was isolated from strain S288C and digested with: lanes a, *Eco*RI; b, *Eco*RI/*Hin*dIII; and c, *Hin*dIII, respectively. A *Hin*dIII restriction digest of phage lambda DNA was included for molecular weight standards. Following transfer to nitrocellulose (Southern, 1975), the filters were probed with radioactively labeled mpHS.3, a single-stranded M13 clone containing the 1·0 kb *Hin*dIII–*Sal*I fragment of the *SUF12*$^+$ coding region (Fig. 3). The filters were then washed with moderate stringency (2 × SSC, 65 °C). The minor bands in these lanes were lost in high stringency washes (0·1 × SSC, 65 °C; data not shown). Southern blot analysis of genomic DNA probed with plasmids containing the *TEF1*$^+$, *TEF2*$^+$ and *SUF12*$^+$ genes showed that the minor bands detected in lanes a to c did not co-migrate with restriction fragments generated from the *TEF*$^+$ genes (unpublished results). Gene disruptions: the black arrow depicts the open reading frame of the *SUF12*$^+$ gene (Fig. 3). Genomic DNA was isolated from the following strains and digested with *Sal*I: lanes d, PW145; e, PW145[*suf12 : : URA3*$^+$]; f and g, 2 spore colonies from a representative PW145[*suf12 : : URA3*$^+$] tetrad. The filters were probed with radioactively labeled mpHS.3 and washed with high stringency (0·1 × SSC, 65 °C). The 4·5 kb fragment is the size of the wild-type *Sal*I fragment predicted by the restriction map of pPW12.1. The 5·3 kb fragment in the heterozygote (lane 2e) reflects replacement of the 0·33 kb *Hin*dIII fragment of the *SUF12*$^+$ coding region with a 1·1 kb fragment containing the *URA3*$^+$ gene. The presence of the 4·5 kb band but not the 5·3 kb band in lanes f and g indicates that *suf12-null* alleles are recessive lethals. The position of the *LEU2*$^+$ gene within *SUF12*$^+$ in PW145[*suf12 : : LEU2*$^+$] is also shown diagrammatically, below. H, *Hin*dIII; Hp, *Hpa*I; S, *Sal*I.

and heterozygous for several markers tightly linked to *suf12* on chromosome IV (Table 1). The growth and sporulation efficiency of *suf12 : : URA3*$^+$/*SUF12*$^+$ heterozygotes are comparable to those of the parental diploid. Each tetrad from several independent Ura$^+$ diploids produced two non-viable, ungerminated spores and two Ura$^-$ colonies. The lethal phenotype was genetically linked to *suf12 : : URA3*$^+$ and *hom2* (data not shown). Southern blot analysis of DNA isolated from the parental diploid, a representative *suf12 : : URA*$^+$/*SUF12*$^+$ heterozygous diploid, and several viable spore colonies indicated that both the wild-type and null alleles were present in the heterozygote, but only wild-type alleles were recovered in tetrads (Fig. 2 lanes d to g).

We then asked if lethality is simply due to an effect on spore germination or to loss of an essential function in vegetative growth. A null allele was constructed by replacing the *Sal*I–*Hpa*I and *Hpa*I–*Hpa*I fragments of the *SUF12*$^+$-coding region with the *LEU2*$^+$ gene (Fig. 2) and transplaced into an appropriately marked diploid (PW145, Table 1). Correct transplacement was inferred from tetrad analysis showing genetic linkage of spore lethality to *suf12 : : LEU2*$^+$ and *hom2* (data not shown). To rescue *suf12-null* allele with an extrachromosomal *SUF12*$^+$ gene, a *suf12 : : LEU2*$^+$/*SUF12*$^+$ heterozygote was transformed with pPWE.RX, a high copy number plasmid containing the yeast 2 μm origin of replication, the *URA3*$^+$ gene and *SUF12*$^+$. Following sporulation, tetrads segregated predominantly 4 : 0 for Ura$^+$ and 2 : 2 for Leu$^+$, but Leu$^+$ Ura$^-$ colonies were not observed. These results indicate that the *suf12 : : LEU2*$^+$ allele is maintained by pPWE.RX and that increased gene dosage of the *SUF12*$^+$ gene is tolerated in these haploids. During repeated subculturing of Leu$^+$ Ura$^+$ spore colonies in complete non-selective media, we did not observe *suf12 : : LEU2*$^+$ colonies in the absence of an extrachromosomal *SUF12*$^+$ gene. To ensure that stable association of the Leu$^+$ and Ura$^+$ phenotypes was not due to plasmid integration into the genome, single colony isolates were mated to an appropriately marked strain (PY39, Table 1) and individual zygotes were selected by micromanipulation. Since Leu$^+$ Ura$^-$ colonies segregated from these *Ura*$^+$ *Leu*$^+$ diploids during vegetative growth, we attribute lethality of the *suf12-null* allele to loss of an essential vegetative function.

(d) SUP35$^+$ *and* SUF12$^+$ *are identical genes*

We have utilized pPWE.RX and the deletion derivatives of pPW12.1 to determine if *SUP35*$^+$ and *SUF12*$^+$ are identical genes. A strain containing *ura3-52* and *sup35-4*, a recessive allele that confers temperature-sensitive growth (Song & Liebman, 1987; Table 1), was transformed with pPW12.1, pPWE.RX and each of the *SUF12*$^+$ deletion derivatives. Ura$^+$ transformants were selected at 25 °C and screened for the ability to grow at 37 °C. Both pPW12.1 and pPWE.RX conferred temperature-independent growth whereas the deletion derivatives did not, showing that the *SUF12*$^+$ gene complements the *sup35-4* allele.

An additional test of allelism was provided by phenotypic expression of the recessive *sup35-4* allele in a diploid containing a *suf12-null* allele. Individual zygotes were isolated by micromanipulation from a mating between SL797-2C and a strain containing a *suf12-null* allele maintained by an extrachromosomal copy of *SUF12*$^+$ (PW145-3C[pPWE.RX], Table 1). During subculturing in complete non-selective media, the recessive *sup35-4* mutation conferred temperature-sensitive growth to those diploids that had lost pPWE.RX. Since diploids heterozygous for a *suf12-null* allele are not temperature-sensitive for growth, expression of the recessive *sup35-4* allele is due to disruption of the allelic wild-type gene, *SUF12*$^+$.

(e) *DNA sequence analysis*

Portions of pPW12.1 were subcloned and the DNA sequence was determined (see Materials and Methods). The *SUF12*$^+$ gene was identified as a 2055 bp open reading frame containing a *Sal*I, *Hin*dIII and *Bst*EII deletion endpoint of pPW12.SΔ, pPW12.HΔ and pPW12.BsΔ, respectively (Fig. 3). The sequence predicts an 88×10^3 M_r protein of 685 amino acids if translation initiates at the first AUG in the open reading frame.

Consensus signals for mRNA slicing (Langford & Gallwitz, 1983) were not observed within the open reading frame or flanking sequences. Three TATAA-like elements are located at −202 (TATTATA), −182 (TATAATAT), and −104 (TATATT). The consensus transcription termination signal proposed by Zaret & Sherman (1982) of TAG...TA(T)GT...TTT occurs at position +2168. The termination/polyadenylation signal sequence TAAATAAG proposed by Bennetzen & Hall (1982*a*) and the termination signal TTTTTATA proposed by Henikoff *et al.* (1983) are not present.

Many components of the yeast translational machinery appear to be co-ordinately regulated, in part, at the level of transcription (Gorenstein & Warner, 1976; Himmelfarb *et al.*, 1985; Warner *et al.*, 1985; Donovan & Pearson, 1986). Transcriptional activation is mediated through specific binding of a protein factor (TUF) to the upstream activating sequences HOMOL1 and/or the RPG box (Teem *et al.*, 1984; Huet *et al.*, 1985; Leer *et al.*, 1985; Rotenberg & Woolford, 1986; Woudt *et al.*, 1986). These promoter elements have been found 50 to 500 bp upstream from the transcriptional start site of most of the ribosomal protein genes (Teem *et al.*, 1984; Leer *et al.*, 1985), the genes encoding the elongation factor EF-1α (Huet *et al.*, 1985), and *SUP45*$^+$ (Breining & Piepersberg, 1986). We were unable to identify a convincing example of either element within the 500 bp 5′ to the TATAA-like elements (Fig. 3; and unpublished results). Since the

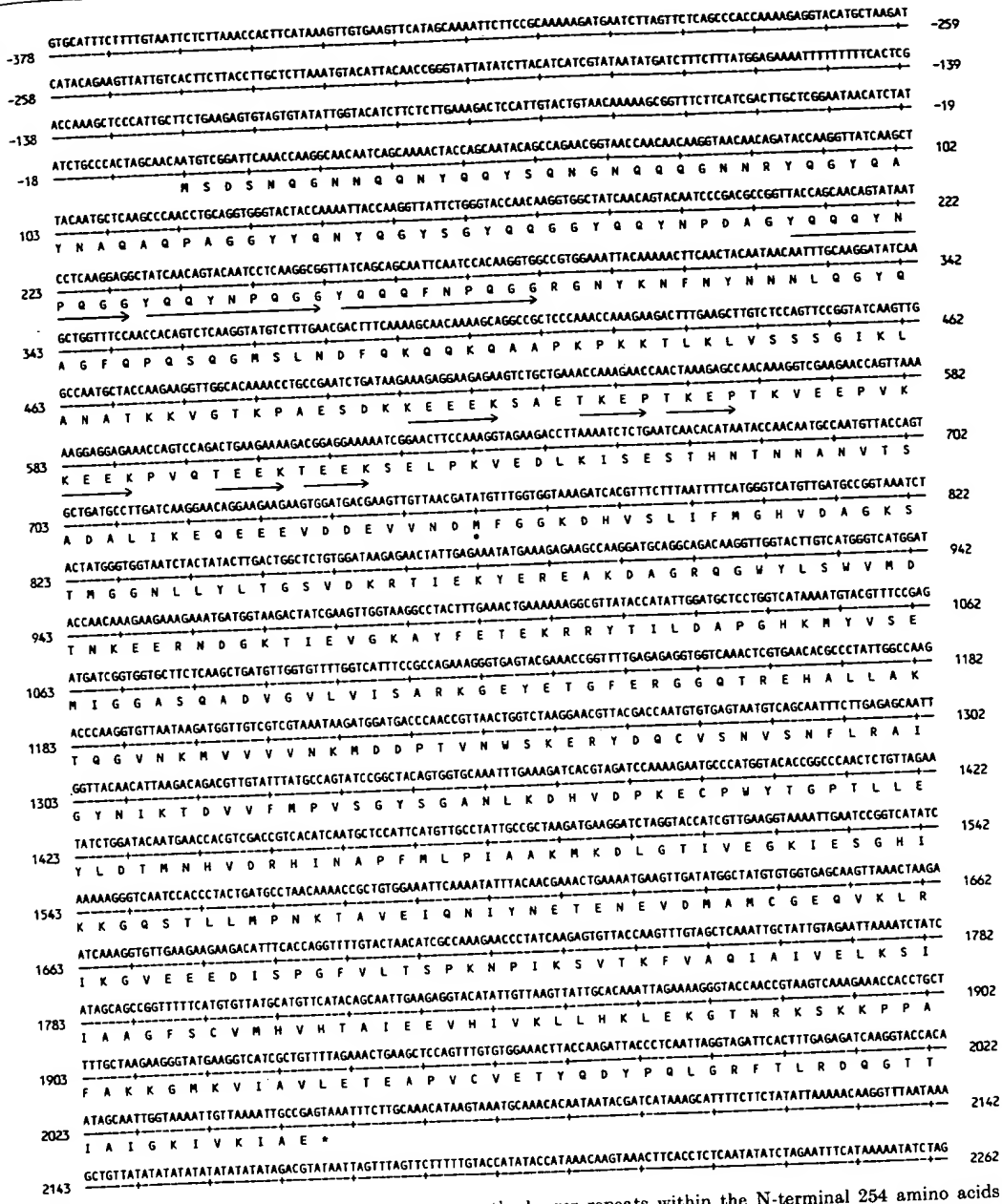


Figure 3. *SUF12*$^+$ DNA and amino acid sequence: the longer repeats within the N-terminal 254 amino acids are indicated by bars. This region also contains several short repeats of 3 to 4 amino acids. The position of Met254 that initiates the SUF12/EF-1α homology is marked (●).

precise sequence parameters of these regulatory sites are not well defined (Huet *et al.*, 1985; Woudt *et al.*, 1986), the lack of HOMOL1/RPG sequence homology may not reflect a lack of co-ordinate transcriptional control. In addition, it is not clear that such control is mediated exclusively through these activation sites (Vignais *et al.*, 1987).

The pattern of codon usage in the *SUF12*$^+$ gene is similar to genes that are not highly expressed (Table 3). A codon bias index of 0·40 also suggests that *SUF12*$^+$ is probably not an abundant protein (Bennetzen & Hall, 1982*b*). The codon bias index reflects the degree of preferred codon usage and has been correlated with the level of mRNA abundance (Bennetzen & Hall, 1982*b*). These observations are consistent with the apparent *SUF12*$^+$ transcript abundance (Fig. 1) that is similar to that of the average-abundance *URA3*$^+$ transcript (Bach *et al.*, 1979).

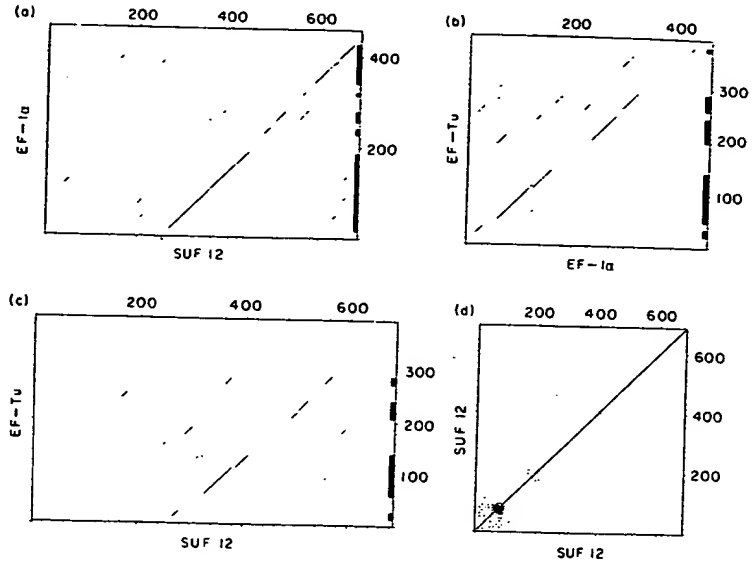


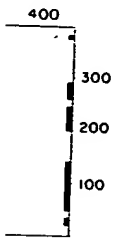
Figure 4. Dot-matrix comparisons between SUF12 and members of the EF-1 family of soluble factors: each dot in (a) SUF12/yeast EF-1α (b) yeast EF-1α/*E. coli* EF-Tu and (c) SUF12/*E. coli* EF-Tu represents 8 exact matches in a window of 30 amino acids. Homology between the C-terminal 30 amino acids of 2 sequences is not detected with this program (Compare, University of Wisconsin Genetics Computer Group). The black bars on the vertical axis compare the relative position of the homology among the proteins. (d). A comparison of the SUF12 sequence with itself showing the extent and position of the repeated elements. Each dot represents 4 exact matches in a window of 6 amino acids. The randomized versions of the N-terminal segments described in the text were obtained with the Shuffle computer program (University of Wisconsin Computer Group).

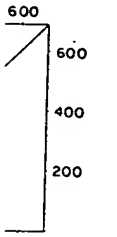
Table 3

Amino acid composition and codon usage

No.		SUF12	Y-L	Y-H	No.		SUF12	Y-L	Y-H	No.		SUF12	Y-L	Y-H
Gly 60	GGG	0·030	0·095	0·007	Asp 30	GAU	0·700	0·655	0·420	Val 50	GUG	0·100	0·180	0·032
	GGA	0·050	0·167	0·005		GAC	0·300	0·345	0·580		GUA	0·180	0·187	0·010
	GGU	0·750	0·557	0·950							GUU	0·520	0·402	0·545
	GGC	0·170	0·180	0·037	Asn 45	AAU	0·530	0·560	0·140		GUC	0·200	0·230	0·412
Ala 43	GCG	0·000	0·087	0·010		AAC	0·470	0·440	0·860					
	GCA	0·160	0·275	0·030						Met 19	AUG	1·000	1·000	1·000
	GCU	0·470	0·392	0·680	Glu 57	GAG	0·230	0·285	0·085					
	GCC	0·370	0·245	0·282		GAA	0·770	0·715	0·915	Ile 32	AUA	0·090	0·233	0·020
											AUU	0·530	0·507	0·453
Ser 35	AGU	0·140	0·155	0·028	Gln 53	CAG	0·250	0·290	0·055		AUC	0·380	0·263	0·527
	AGC	0·060	0·100	0·027		CAA	0·770	0·710	0·945					
	UCG	0·090	0·083	0·015						Leu 35	UUG	0·430	0·326	0·750
	UCA	0·170	0·195	0·038	Arg 18	AGG	0·060	0·187	0·023		UUA	0·200	0·272	0·133
	UCU	0·340	0·317	0·528		AGA	0·610	0·500	0·867		CUG	0·090	0·103	0·022
	UCC	0·200	0·155	0·362		CGG	0·000	0·030	0·000		CUA	0·200	0·133	0·070
						CGA	0·000	0·057	0·000		CUU	0·090	0·112	0·022
Thr 39	ACG	0·030	0·120	0·015		CGU	0·330	0·177	0·107		CUC	0·000	0·055	0·003
	ACA	0·210	0·292	0·055		CGC	0·000	0·050	0·002					
	ACU	0·360	0·357	0·485						Trp 4	UGG	1·000	1·000	1·000
	ACC	0·410	0·227	0·445	Lys 66	AAG	0·580	0·445	0·810					
						AAA	0·420	0·555	0·190	Tyr 35	UAU	0·290	0·525	0·130
Pro 30	CCG	0·000	0·097	0·007							UAC	0·710	0·475	0·870
	CCA	0·600	0·452	0·860	His 13	CAU	0·540	0·645	0·260					
	CCU	0·330	0·295	0·125		CAC	0·460	0·355	0·740	Phe 16	UUU	0·560	0·585	0·210
	CCC	0·070	0·155	0·007							UUC	0·440	0·415	0·790
					Cys 5	UGU	0·800	0·705	0·890					
						UGC	0·200	0·295	0·110	* 1	UAA	1·000	0·447	0·870

Synonymous codon usage is the frequency of each codon in the group of synonymous codons for a particular amino acid. The Y-H and Y-L columns were obtained from data presented (see Table 2) by Sharp *et al.* (1986) and correspond to codon usage of group A, 66 highly expressed yeast genes and group B, 38 lowly expressed yeast genes, respectively. We have presented usage as a fraction of one, whereas Sharp *et al.* presented usage as a fraction of the number of synonymous codons for a particular amino acid.





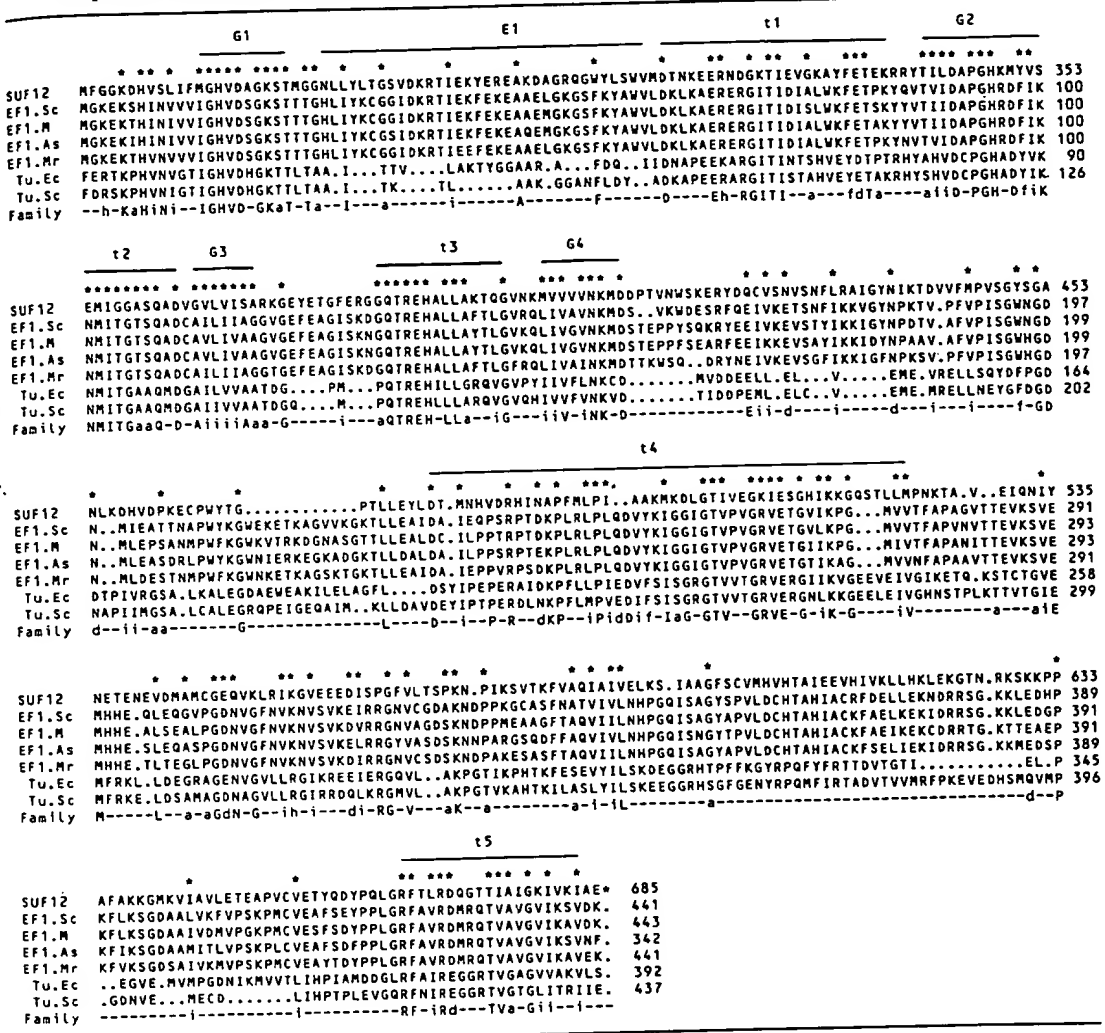
soluble factors: each dot in (a)
s 8 exact matches in a window
ot detected with this program
tical axis compare the relative
with itself showing the extent
ndow of 6 amino acids. The
the Shuffle computer program

	SUF12	Y-L	Y-H
GUG	0·100	0·180	0·032
GUA	0·180	0·187	0·010
GUU	0·520	0·402	0·545
GUC	0·200	0·230	0·412
AUG	1·000	1·000	1·000
AUA	0·090	0·233	0·020
AUU	0·530	0·507	0·453
AUC	0·380	0·263	0·527
UUG	0·430	0·326	0·750
UUA	0·200	0·272	0·133
CUG	0·090	0·103	0·022
CUA	0·200	0·133	0·070
CUU	0·090	0·112	0·022
CUC	0·000	0·055	0·003
UGG	1·000	1·000	1·000
UAU	0·290	0·525	0·130
UAC	0·710	0·475	0·870
UUU	0·560	0·585	0·210
UUC	0·440	0·415	0·790
UAA	1·000	0·447	0·870

particular amino acid. The Y-H
d to codon usage of group A, 66
nted usage as a fraction of one,
r amino acid.

Table 4

Comparison of SUF12 amino acid sequence with a consensus sequence of the EF-1 family of proteins



To localize the residues conserved among members of the EF-1 family and SUF12, the SUF12 sequence was aligned with a consensus sequence (Family) derived from the members of the EF-1 family presented in this Table. The consensus sequence represents a stringent basis of comparison, since an identical or conservative amino acid substitution must be present in each of the factors to be included in the consensus sequence. Alignments were first generated by the Gap and Pretty computer programs (University of Wisconsin Computer Group) and then adjusted by eye to maximize homology within the EF-1 family. In the consensus sequence exact amino acid matches are in upper case and conservative Dayhoff amino acid substitutions (Dayhoff, 1978) are in lower case. The following Dayhoff conservative categories were used: a = S,A,P,T,G; d = D,E,Q,N; h = K,R,H; i = V,L,I,M; f = Y,W,F; c = C). The asterisks above the SUF12 sequence represent either an exact match or a conservative substitution. The regions of SUF12 homology to GTP binding proteins (G1 to G4), conserved regions of the EF-1 family (t1 to t5), and a region characteristic of the eukaryotic group (E1) are indicated. SUF12, the region of SUF12 homology to the EF-1 family. EF1.Sc, EF-1α from yeast (Nagata *et al.*, 1984); EF1.M, EF-1α from mammals (Brands *et al.*, 1986); EF1.As, EF-1α from *Artemia salina* (van Hemert *et al.*, 1984); EF1.Mr, EF-1α from *Mucor racemosus* (Linz *et al.*, 1986): Tu. Ec, EF-Tu from *E. coli*. (An & Friesen, 1982): Tu,Sc, EF-Tu from yeast mitochondria (Nagata *et al.*, 1983).

(f) *Analysis of the inferred SUF12 protein*

In a search for similar amino acid sequences in the National Biomedical Research Foundation (NBRF) data bank of protein sequences (George *et al.*, 1986), we found significant homology to the EF-1 family of elongation factors. Pairwise comparisons between SUF12, yeast EF-1α and *E. coli* EF-Tu (Fig. 4(a) to (c)) illustrate the following observations: (1) homology between SUF12 and members of the EF-1 family begins internally at amino acid 254 of SUF12. (2) Homology is more extensive between SUF12 and EF-1α (36% exact matches; 65% with conservative substitutions) than between

EF-1α and EF-Tu of *E. coli* (28% identity; 50% with conservative substitutions). (3) SUF12/EF-Tu homology is less extensive (20% identity; 50% conservative substitutions) but located in similar positions to EF-1α/EF-Tu and SUF12/EF-1α homology.

By comparison of the SUF12 amino acid sequence with a consensus sequence of the EF-1 family of proteins (Table 4), we found that SUF12 contains 81% of the conserved residues of this family if conservative substitutions are permitted. We have identified three groups of sequence homology in the SUF12 protein that exhibit the most extensive identity to either the consensus sequence of the EF-1 family or, more specifically, to the eukaryotic factors. Regions G1 to G4 contain residues conserved among GTP-binding proteins. Regions t1 to t5 of SUF12 are similar to regions that have been conserved in both prokaryotic and eukaryotic members of the EF-1 family, while E1 appears to be a characteristic of eukaryotic EF-1α proteins.

Regions G1 to G4 contain conserved features of GTP-binding proteins. These regions have been associated with GTP binding and hydrolysis in EF-Tu (Kaziro, 1978), and exhibit the most extensive homology between SUF12 and the EF-1 family (Table 4). Homologous regions have been identified in other GTP-binding proteins, including several translation factors (Sacerdot *et al.*, 1984; Zengel *et al.*, 1984; Kohno *et al.*, 1986), the p21 family of *ras* oncogene proteins (Halliday, 1983) and bovine transducin (Lochrie *et al.*, 1985). X-ray diffraction analysis of a trypsin-modified form of EF-Tu·GDP indicates that these regions are in the vicinity of the GDP-binding site, and most of the invariant amino acids interact directly with the GDP ligand (Jurnak, 1985; la Cour *et al.*, 1985). The most conserved elements among GTP-binding proteins correspond to Gly267-His-Val-Asp-Ser-Gly-Lys273 and Asn406-Lys-Met-Asp409 of SUF12.

Regions t1 to t5 of SUF12 correspond to regions that have been conserved in the EF-1 family (Table 4). Residues of *E. coli* EF-Tu that have been implicated in tRNA binding include His66 (t1), His118 (t3) (Jonak *et al.*, 1984), Lys208 and Lys237 (t4) van Noort *et al.*, 1984, 1985) and Cys81 (G2) (Miller *et al.*, 1971; Arai *et al.*, 1974; Jonak *et al.*, 1982). The t1 region is also similar to a segment of mammalian elongation factor 2 (EF-2) and *E. coli* elongation factor G (EF-G). Kohno *et al.* (1986) have proposed that this region of homology may reflect a common function of elongation factors other than GTP binding. The functional significance of the conserved t2 and t5 regions is not known.

The E1 region of SUF12 is homologous to a segment that is highly conserved among the eukaryotic members of the EF-1 family (90% identity; 97% with conservative substitutions). Although other regions of the eukaryotic EF-1α proteins also lack a prokaryotic counterpart (Brands *et al.*, 1986), E1 is unusual in its length and position in the most conserved region of this family. A segment of similar conserved character and/or sequence is not present in this position in other soluble factors or GTP-binding proteins (Table 4; Brands *et al.*, 1986; Kohno *et al.*, 1986).

SUF12/EF-1α homology is essentially co-linear and located in the C-terminal two-thirds of the SUF12 protein (Fig. 4). The N-terminal 120 amino acid residues exhibit unusual composition (45% Asn and Gln, 33% Gly and Tyr) and extensive repeated structure (Figs 3 and 4(d)) with a strong potential for beta-sheet formation (Garnier *et al.*, 1978; Chou & Fasman, 1978). Allowing conservative substitutions, this segment contains three nearly identical repeats of nine to ten amino acids (Gln-Gln-(Gln)-Tyr(Phe)-Asn-Pro-Gln-Gly-Gly-Phe) and several shorter repeats (Fig. 3). To determine if these repeats result from the limited amino acid composition of this segment (78% Gln, Asn, Tyr and Gly), we compared the linear sequence of the first 120 amino acid residues with a randomized sequence generated from the same residues. Dot-matrix analysis of this comparison indicates that at least the long repeats are not simply due to the amino acid composition. The adjacent segment of 134 contains a high density of charged residues (41% Glu, Asp and Lys). This region contains a tandem repeat of Thr-Lys-Glu-Pro and several dispersed repeats similar to Lys(Thr)-Glu-Glu-(Glu)-Lys (Figs 3 and 4(d)).

4. Discussion

To understand the molecular mechanism of *suf12*-mediated misreading and translational frame-shifting, we have isolated the *SUF12*$^+$ gene by complementation and identified the 2·3 kb *SUF12*$^+$ transcript. *SUF12*$^+$ is an essential single-copy gene that is identical with *SUP35*$^+$ (Hawthorne & Leupold, 1974). DNA sequence analysis has revealed an uninterrupted open reading frame sufficient to encode an 88×10^3 M_r protein of 685 amino acids. The amino acid composition suggests that SUF12 is a soluble protein containing approximately equal numbers of acidic and basic residues (see Table 3).

(a) *Homology between the* SUF12 *protein and the EF-1 family of elongation factors*

The SUF12 amino acid sequence exhibits extensive homology to the EF-1 family of elongation factors (Fig. 4), which includes 81% of the conserved residues of this family (Table 4). Beginning internally at amino acid 254 of SUF12, homology is more extensive between SUF12 and EF-1α of *S. cerevisiae* (36% identity; 65% with conservative substitutions) than between EF-1α and *E. coli* EF-Tu (28% identity; 50% with conservative substitutions). The corresponding DNA homology between *SUF12*$^+$ and *TEF2*$^+$, one of the genes encoding yeast EF-1α, is 50 to 65% depending on the window length and position. The most conserved regions of the EF-1 family are

involved in GTP and aminoacyl-tRNA binding, and the essentially co-linear homology in SUF12 suggests that SUF12 may be functionally related to this family of elongation factors.

The segment of SUF12 between Lys259 and Asp409 exhibits the strongest homology to yeast EF-1α. This segment contains regions G1 to G4 (Table 4), which have been associated with GTP binding and hydrolysis in EF-Tu (Jurnak, 1985; la Cour *et al.*, 1985), and several other GTP-binding proteins (Halliday, 1983; Sacerdot *et al.*, 1984; Zengel *et al.*, 1984; Lochrie *et al.*, 1985; Kohno *et al.*, 1986). In translation, GTP binding and hydrolysis are predominantly associated with the soluble translation factors directly involved in either the translation cycle or ribosome assembly (Kaziro, 1978; Moldave, 1985). GTP/GDP binding appears to induce changes in factor conformation that mediate ribosome binding and release (Kaziro, 1978; Thompson *et al.*, 1986). These observations suggest that SUF12 is a soluble factor, and the apparent guanosine-binding domains of SUF12 may reflect a similar recycling function.

Other regions conserved in the EF-1 family include residues of EF-Tu (His66, Cys81, His118, Lys208, Lys237) that have been associated with the binding site for aminoacyl-tRNA (Miller *et al.*, 1971; Arai *et al.*, 1974; Jonak *et al.*, 1982, 1984; van Noort *et al.*, 1984, 1985). Each of these EF-Tu residues is located within highly conserved regions of the EF-1 family and the corresponding regions of SUF12 (G2, t1, t3 and t4 in Table 4). Of the three residues that are invariant in the EF-1 family, two are conserved in SUF12 and correspond to His389 (t3) and Lys515 (t4) (Table 4). The proximity of tRNA binding domains to the GTP binding and hydrolysis center of EF-Tu (EF-1α) probably reflects the allosteric effects of GTP/GDP on aminoacyl-tRNA binding to EF-Tu (Kaziro, 1978). Although SUF12 is homologous to the functionally conserved regions of the EF-1 family, further experiments are required to determine if SUF12 exhibits either GTP or tRNA binding.

(b) SUF12 *is essential* in vivo *and functionally distinct from EF-1α*

SUF12/EF-1α homology is not as extensive as homology among EF-1 proteins from different eukaryotic species, including yeast (Nagata *et al.*, 1984), brine shrimp (van Hemert *et al.*, 1984), mammals (Brands *et al.*, 1986) and slime mold (Linz *et al.*, 1986). The homology within the eukaryotic group of factors is approximately 75% identity (Table 4), in sharp contrast to the observed 36% SUF12/EF-1α identity. This observation suggests that, although SUF12 and EF-1α share structural features and possibly functional activities, these proteins are probably not functionally equivalent.

A functional distinction between SUF12 and EF-1α is supported by an *in-vivo* requirement for an intact *SUF12*$^+$ gene. We were unable to recover *suf12-null* alleles from sporulated diploids (Fig. 2 lanes d to g) or vegetative haploids without rescue by an extrachromosomal *SUF12*$^+$ gene. The lethality associated with *suf12-null* alleles indicates that EF-1α cannot compensate for loss of SUF12 function. Likewise, the SUF12 protein does not compensate for loss of EF-1α activity, since at least one functional EF-1α-encoding gene, *TEF1*$^+$ or *TEF2*$^+$, is also necessary for viability (Cottrelle *et al.*, 1985).

The linear SUF12 sequence suggests that SUF12 has evolved as a fusion protein of unique N-terminal SUF12 domains fused to domains structurally related to EF-1α (Fig. 4(a)). SUF12 may reflect a gene duplication event and subsequent divergence of function from EF-1α. The N-terminal 120 amino acid residue segment of SUF12 appears to be soluble due to a high concentration (45%) of soluble acid amines (Asn and Gln) and lack of hydrophobic residues. Analyses of secondary structure (Chou & Fasman, 1978; Garnier *et al.*, 1978) indicate a strong potential for beta-sheet formation. Allowing conservative amino acid substitutions, this region contains three repeats of nine amino acids and several shorter repeating units (Fig. 3). The adjacent segment of 134 amino acids contains a high density (41%) of charged residues (Lys, Glu and Asp) and also several short repeats (Fig. 3). We were unable to identify significant sequence homology between the N-terminal 254 amino acid residues of SUF12 and other proteins in the NBRF data bank (George *et al.*, 1986).

(c) *Comparison of SUF12 with other GTP-dependent translation factors*

Since GTP binding and hydrolysis are characteristics of several translation factors, we considered the possibility that SUF12 may correspond to a soluble factor that has been identified. On the basis of direct or inferred amino acid sequence analysis and/or molecular weight comparisons, SUF12 does not correspond to elongation factor 2 (EF-2), a GTP-dependent factor that mediates translocation (van Ness *et al.*, 1978) or the 33×10^3 M_r alpha subunit of initiation factor 2 (IF-2), a factor involved in delivery of the initiator tRNA to the preinitiation complex (Baan *et al.*, 1981). The DNA restriction maps of *SUF12*$^+$ and the gene encoding elongation factor 3 (EF-3) are not related (Qin *et al.*, 1987). EF-3 is a soluble GTP-dependent factor (125×10^3 M_r) required in translation with yeast ribosomes (Skogerson & Wakatama, 1976; Skogerson & Engelhardt, 1977; Dasmahapatra & Chakraburtty, 1981). Cytoplasmic factors with molecular weights substantially different from that of SUF12 (88×10^3) include factors that co-purify with EF-1α, EF-1β (33×10^3) and EF-1γ (47×10^3) (Saha & Chakraburtty, 1986).

Our analysis suggests that SUF12 is not a yeast release factor. Termination in eukaryotes appears to be GTP-dependent and mediated by a single release factor that recognizes each of the three stop codons (Caskey, 1980). The molecular weight

(56×10^3) of the release factor isolated from rabbit reticulocytes (Caskey, 1980) is substantially smaller than the predicted molecular weight (88×10^3) of SUF12. The *E. coli* release factors RF1 and RF2 (Craigen *et al.*, 1985) do not exhibit significant homology to either EF-Tu or SUF12 (data not shown). In *E. coli*, increased levels of RF1 and RF2 enhance termination in a competition between release factors and suppressor tRNAs for recognition of nonsense codons (Caskey *et al.*, 1984; Weiss *et al.*, 1984). We have preliminary results indicating that *SUF12*$^+$ does not exhibit a similar phenotype when overexpressed in yeast (unpublished results). In addition, misreading of poly(U) templates in *sup35* cell-free systems (Eustice *et al.*, 1986) is difficult to reconcile with a defect in termination.

(d) *Co-dominant* suf12 *expression*

We interpret *suf12* allele-specific suppression as evidence that *suf12*-mediated misreading is not simply a loss or reduction of wild-type function. This view is supported by our analysis of *suf12/SUF12*$^+$ and *suf12-null/SUF12*$^+$ heterozygous diploids indicating co-dominant effects of *suf12* mutations. Although *suf12*-mediated suppression is phenotypically recessive, the presence of a wild-type *SUF12*$^+$ gene does not mask the poor sporulation efficiency and spore viability characteristics of these strains (Hawthorne & Leupold, 1974; Culbertson *et al.*, 1982). In *suf12/SUF12*$^+$ heterozygous crosses, poor viability is not limited to spores containing *suf12* alleles, but affects *SUF12*$^+$ spores as well. The frequency of inviable spores varies with the *suf12* allele involved but typically ranges from 20 to 40% in heterozygous crosses.

If the *suf12* sporulation and suppressor phenotypes resulted from variation in residual wild-type function imposed by distinct *suf12* mutations, we would expect diploids containing a *suf12-null* allele to exhibit a more severe phenotype than the phenotype associated with *suf12* suppressor mutations. For example, a large proportion of tetrads might segregate 0 : 4 and 1 : 3 for spore viability. However, this is not the case, since *suf12-null/SUF12*$^+$ diploids sporulate efficiently and produce two viable spores in virtually every tetrad. Also, we have not detected nonsense or frameshift suppression in *suf12-null/SUF12*$^+$ diploids, as might be expected if suppression resulted from reduced wild-type activity. These results indicate co-dominant *suf12* expression and suggest that allele-specific suppression may result from functionally distinct *suf12* mutant proteins.

(e) *SUF12 function in translation*

To construct a model of SUF12 function, our molecular and genetic analysis must be integrated with the biochemical behavior of *suf12* mutants *in vitro*. Previous studies utilizing cytoplasmic ribosomes isolated from *sup35* (Eustice *et al.*, 1986) and *sup2* (Surguchov *et al.*, 1984) strains suggest that *suf12*-mediated misreading is associated with the ribosome and not a soluble component of the *in-vitro* system. As a result of these observations, the *SUF12*$^+$ gene product has been postulated to be an enzyme involved in ribosome modification or a component of the ribosome. Attempts to identify the defective subunit and the specific ribosomal proteins involved have led to conflicting results (Surguchov *et al.*, 1984; Eustice *et al.*, 1986). A modification activity is difficult to reconcile with allele-specific, co-dominant expression of *suf12* mutations and SUF12 sequence homology to EF-1α. On the other hand, SUF12 does not appear to be a ribosomal protein, since the predicted molecular weight (88×10^3) is almost twice as large as any known ribosomal protein in yeast (Warner, 1982). Furthermore, the pattern of codon usage (Table 3) and codon bias index (0·40) suggests that *SUF12*$^+$ is not a highly expressed gene, in contrast to the genes encoding ribosomal proteins (Bennetzen & Hall, 1982*b*; Sharp *et al.*, 1986).

An alternative explanation for an apparent ribosomal defect in *suf12(sup35)* mutant extracts involves an association of SUF12 with ribosomes or polysomes in cell-free systems that is lost in preparations of ribosomal proteins. This interpretation is supported by similar behavior of EF-3, an essential translation factor required by yeast ribosomes but not ribosomes isolated from other eukaryotes (Skogerson & Wakatama, 1976; Skogerson & Engelhardt, 1977). EF-3 is found predominantly associated with polysomes but apparently not with the ribosomal subunits (Hutchison *et al.*, 1984). Furthermore, EF-3 does not appear to be tightly bound to polysomes, since it is lost by centrifugation in low-salt sucrose gradients (Hutchison *et al.*, 1984). EF-3 is also similar to SUF12 in that both of these proteins are probably not present in amounts stoichiometric with the ribosome (Hutchison *et al.*, 1984).

(f) *The relationship between* suf12*-mediated translational frameshifting and codon recognition errors*

We propose that *suf12*-mediated translational frameshifting is a consequence of a defect in the normal process of codon recognition. Experiments in *E. coli* indicate that translational frameshifting may occur *in vivo* and *in vitro* as a result of competition among tRNAs (Atkins *et al.*, 1979; Weiss & Gallant, 1983, 1986), and possibly between tRNAs and release factors (Craigen *et al.*, 1985). One interpretation of these results is that competing tRNAs bind out-of-frame relative to the codon that occupies the ribosomal A site. This type of translational frameshifting is consistent with the observed doublet decoding of Ala and Pro codons by *E. coli* $tRNA^{Ser3}$ and $tRNA^{Thr4}$, respectively, that involves conventional base-pairing at two positions of the codon : anticodon pair (Bruce *et al.*, 1986; Dayhuff *et al.*, 1986). We infer from these

et al., 1984) strains suggest misreading is associated with a soluble component of the in- ult of these observations, the t has been postulated to be an ribosome modification or a iosome. Attempts to identify ; and the specific ribosomal ve led to conflicting results 1984; Eustice *et al.*, 1986). y is difficult to reconcile with ninant expression of *suf12* 2 sequence homology to EF- 1, SUF12 does not appear to otein, since the predicted $< 10^3$) is almost twice as large nal protein in yeast (Warner, the pattern of codon usage ias index (0·40) suggests that ly expressed gene, in contrast oding ribosomal proteins 82*b*; Sharp *et al.*, 1986). planation for an apparent *uf12(sup35)* mutant extracts 1 of SUF12 with ribosomes or e systems that is lost in mal proteins. This interpreta- similar behavior of EF-3, an factor required by yeast bosomes isolated from other on & Wakatama, 1976; ardt, 1977). EF-3 is found ated with polysomes but h the ribosomal subunits 4). Furthermore, EF-3 does ly bound to polysomes, since ugation in low-salt sucrose *et al.*, 1984). EF-3 is also iat both of these proteins are in amounts stoichiometric tchison *et al.*, 1984).

p between suf12*-mediated ameshifting and codon nition errors*

suf12-mediated translational sequence of a defect in the on recognition. Experiments t translational frameshifting nd *in vitro* as a result of RNAs (Atkins *et al.*, 1979; 1986), and possibly between ctors (Craigen *et al.*, 1985). these results is that compet- f-frame relative to the codon osomal A site. This type of fting is consistent with the ding of Ala and Pro codons nd $tRNA^{Thr4}$, respectively, tional base-pairing at two anticodon pair (Bruce *et al.*, 1986). We infer from these observations that one type of translational error may result from a failure of the translational machinery to assess the position of tRNA binding relative to the ribosomal A site and/or the number of bases involved in the codon–anticodon interaction.

The Internal Standard Hypothesis (Thompson *et al.*, 1986) asserts that in *E. coli* discrimination between cognate and near-cognate tRNA species resides in the dissociation rates of these species from the ribosome relative to an internal standard, the dissociation rate of the EF-Tu·GDP complex. A reasonable application and extension of the Thompson model incorporates the kinetics of out-of-frame tRNA binding and dissociation. This mechanism does not limit the position of tRNA binding on the mRNA to the triplet occupying the ribosomal A site or the number of bases involved in a specific tRNA–mRNA interaction to three bases. Our application of the model proposed by Thompson *et al.* (1986) requires only sufficient tRNA–mRNA and/or tRNA–ribosome stability to achieve a slower rate of ribosome dissociation than that of EF-Tu·GDP. Factors that influence the relative rates of EF-Tu·GDP and tRNA dissociation may influence the apparent level of codon misreading (Thompson *et al.*, 1986) as well as the level of translational frameshifting.

Our application of the Internal Standard Hypothesis to translational frameshifting is consistent with the behavior of several translational suppressors in *E. coli* that increase codon misreading and also increase the level of translational frameshifting. These suppressors include the S4 *ram* mutants (Atkins *et al.*, 1972), the *supK* tRNA methylase mutant (Atkins & Ryce, 1974), the *ksgA* ribosomal RNA modification mutant (van Buul *et al.*, 1984), and possibly certain EF-Tu mutants (Hughs, 1986). Conversely, the str^r mutations in the S12 ribosomal protein increase translational accuracy and decrease the level of translational frameshifting (Atkins *et al.*, 1972). Since translational frameshifting is not a standard assay for misreading, we suggest that further analysis of other translational suppressors will support the correlation between codon recognition errors and translational frameshifting.

On the basis of the molecular and genetic analyses of other translational frameshift suppressors, we propose that translational frameshifting in *suf12* mutant strains is a consequence of a defect in codon recognition. We favor a molecular model of translational frameshifting that is dependent on the kinetics of out-of-frame tRNA binding; however, alternative models have been proposed (Kurland, 1980; Weiss, 1983). Our model of translational frameshifting is not intended to explain the high levels of frameshifting that have been observed in the absence of suppressor mutations. For example, translation of the full-length mRNA encoding RF2 of *E. coli* requires a frameshift event to avoid termination at a premature stop codon within the open reading frame (Craigen *et al.*, 1985). Frameshifting at the premature stop codon occurs at a very high rate of 50% (Craigen & Caskey, 1986). Such rates are difficult to reconcile with a model of translational frameshifting that is dependent on errors in codon recognition, suggesting that other mechanisms are operating to achieve these rates (Craigen & Caskey, 1987).

This research was supported by the College of Agricultural and Life Sciences, University of Wisconsin, Madison; National Institutes of Health grant GM26217 (M.R.C.). P.G.W. was supported by Agrigenetics-Lubrizol Fund 133G212 and Public Health Service grant GM07133 in the Graduate Training Program in the Department of Genetics. This is Laboratory of Genetics paper no. 2918.

References

An, G. & Friesen, J. D. (1980). *Gene*, **12**, 33–39.
Arai, K.-I., Kawakita, M., Nakamura, S., Ishikawa, I. & Kaziro, Y. (1974). *J. Biochem.* **76**, 523–534.
Atkins, J. F. & Ryce, S. (1974). *Nature (London)*, **249**, 527–530.
Atkins, J. F., Elseviers, D. & Gorini, L. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 1192–1195.
Atkins, J. F., Gesteland, R. F., Reid, B. R. & Anderson, C. W. (1979). *Cell*, **18**, 1119–1131.
Bach, M.-L., Lacroute, F. & Botstein, D. (1979). *Proc. Nat. Acad. Sci., U.S.A.* **76**, 386–390.
Baan, R. A., Keller, P. B. & Dahlberg, A. E. (1981). *J. Biol. Chem.* **256**, 1063–1066.
Biggen, M. D., Gibson, T. J. & Hong, G. F. (1983). *Proc. Nat. Acad. Sci., U.S.A.* **80**, 3963–3965.
Bennetzen, J. L. & Hall, B. D. (1982*a*). *J. Biol. Chem.* **257**, 3018–3025.
Bennetzen, J. L. & Hall, B. D. (1982*b*). *J. Biol. Chem.* **257**, 3026–3031.
Botstein, D., Falco, S. C., Stewart, S. E., Brennan, M., Scherer, S., Stinchcomb, D. T., Struhl, K. & Davis, R. W. (1979). *Gene*, **8**, 17–24.
Brands, J. H. G. M., Maassen, J. A., van Hemert, F. J., Amons, R. & Moller, W. (1986). *Eur. J. Biochem.* **155**, 167–171.
Breining, P. & Piepersberg, W. (1986). *Nucl. Acids Res.* **14**, 5187–5197.
Broach, J. R., Strathern, J. N. & Hicks, J. B. (1979). *Gene*, **8**, 121–133.
Bruce, A. G., Atkins, J. F. & Gesteland, R. F. (1986). *Proc. Nat. Acad. Sci., U.S.A.* **83**, 5062–5066.
Caskey, C. T. (1980). *Trends Biochem. Sci.* **5**, 234–237.
Caskey, C. T., Forrester, W. C., Tate, W. & Ward, C. D. (1984). *J. Bacteriol.* **158**, 365–368.
Changchien, L.-M., Schwarzbauer, J., Cantrell, M. & Craven, G. R. (1978). *Nucl. Acids Res.* **5**, 2789–2799.
Chou, P. Y. & Fasman, G. D. (1978). *Advan. Enzymol.* **47**, 45–148.
Cottrelle, P., Cool, M., Thuriaux, P., Price, V. L., Thiele, D., Buhler, J.-M. & Fromageot, P. (1985). *Curr. Genet.* **9**, 693–697.
Craigen, W. J. & Caskey, C. T. (1986). *Nature (London)*, **322**, 273–275.
Craigen, W. J. & Caskey, C. T. (1987). *Cell*, **50**, 1–2.
Craigen, W. J., Cook, R. G., Tate, W. P. & Caskey, C. T. (1985). *Proc. Nat. Acad. Sci., U.S.A.* **82**, 3616–3620.
Culbertson, M. R., Gaber, R. F. & Cummins, C. M. (1982). *Genetics*, **102**, 361–378.

Curran, J. F. & Yarus, M. (1986). *Proc. Nat. Acad. Sci., U.S.A.* **83**, 6538–6542.
Dasmahapatra, B. & Chakraburtty, K. (1981). *J. Biol. Chem.* **256**, 9999–10004.
Dayhoff, M. O. (1978). *Atlas of Protein Sequence and Structure*, vol. 5, supp. 1 to 3, National Biomedical Research Foundation, Washington, DC.
Dayhuff, T. J., Atkins, J. F. & Gesteland, R. F. (1986). *J. Biol. Chem.* **261**, 7491–7500.
Donovan, D. M. & Pearson, N. J. (1986). *Mol. Cell. Biol.* **6**, 2429–2435.
Eustice, D. C., Wakem, L. P., Wilhelm, J. M. & Sherman, F. (1986). *J. Mol. Biol.* **188**, 207–214.
Gaber, R. F. & Culbertson, M. R. (1982). *Gene*, **19**, 163–172.
Gaber, R. F. & Culbertson, M. R. (1984). *Mol. Cell. Biol.* **4**, 2052–2061.
Garnier, J., Osguthorpe, D. J. & Robson, B. (1978). *J. Mol. Biol.* **120**, 97–120.
George, D. G., Barker, W. C. & Hunt, L. T. (1986). *Nucl. Acids Res.* **14**, 11–15.
Gerlach, W. L. (1975). *Mol. Gen. Genet.* **138**, 53–63.
Gorenstein, C. & Warner, J. R. (1976). *Proc. Nat. Acad. Sci., U.S.A.* **73**, 1547–1551.
Halliday, K. R. (1983). *J. Cyclic Nucl. Prot. Phosphoryl. Res.* **9**, 435–448.
Hawthorne, D. C. & Leupold, U. (1974). *Curr. Top. Microbiol. Immunol.* **64**, 1–47.
Henikoff, S., Kelly, J. D. & Cohen, E. H. (1983). *Cell*, **33**, 607–614.
Himmelfarb, H. J., Maicas, E. & Friesen, J. D. (1985). *Mol. Cell. Biol.* **5**, 816–822.
Hinnen, A., Hicks, J. B. & Fink, G. R. (1978). *Proc. Nat. Acad. Sci., U.S.A.* **75**, 1929–1933.
Holmes, D. S. & Quigley, M. (1981). *Anal. Biochem.* **114**, 193–197.
Huet, J., Cottrelle, P., Cool, M., Vignais, M.-L., Thiele, D., Marck, C., Buhler, J.-M., Sentenac, A. & Fromageot, P. (1985). *EMBO J.* **4**, 3539–3547.
Hughs, D. (1986). *Mol. Gen. Genet.* **202**, 108–111.
Hutchison, J. S., Feinberg, B., Rothwell, T. C. & Moldave, K. (1984). *Biochemistry*, **23**, 3055–3064.
Ito, H., Fukuda, Y., Murata, K. & Kimura, A. (1983). *J. Bacteriol.* **153**, 163–168.
Jonak, J., Petersen, T. E., Clark, B. F. C. & Rychlik, I. (1982). *FEBS Letters*, **150**, 485–488.
Jonak, J., Petersen, T. E., Meloun, B. & Rychlik, I. (1984). *Eur. J. Biochem.* **144**, 295–303.
Jurnak, F. (1985). *Science*, **230**, 32–36.
Kaziro, Y. (1978). *Biochim. Biophys. Acta*, **505**, 95–127.
Kirsebom, L. A. & Isaksson, L. A. (1985). *Proc. Nat. Acad. Sci., U.S.A.* **82**, 717–721.
Kohno, K., Uchida, T., Ohkubo, H., Nakanishi, S., Nakanishi, T., Fukui, T., Ohtsuka, E., Ikehara, M. & Okada, Y. (1986). *Proc. Nat. Acad. Sci., U.S.A.* **83**, 4978–4982.
Kurland, C. G. (1980). In *Ribosomes Structure, Function, and Genetics*. (Chambliss, G. Craven, G. R., Davies, J., Davies, K., Kahan, L. & Nomura, M., eds), pp. 597–614, University Park Press, Baltimore, MD.
la Cour, T. F. M., Nyborg, J., Thirup, S. & Clark, B. F. C. (1985). *EMBO J.* **4**, 2385–2388.
Langford, C. J. & Gallwitz, D. (1983). *Cell*, **33**, 519–527.
Leer, R. J., van Raamsdonk-Duin, M. M. C., Mager, W. H. & Planta, R. J. (1985). *Curr. Genet.* **9**, 273–277.
Lindquist, S. (1981). *Nature (London)*, **293**, 311–314.
Linz, J. E., Lira, L. M. & Sypherd, P. S. (1986). *J. Biol. Chem.* **261**, 15022–15029.
Lipmann, R. (1969). *Science*, **164**, 1024–1031.
Lochrie, M. A., Hurley, J. B. & Simon, M. I. (1985). *Science*, **228**, 96–99.
Mandel, M. & Higa, A. (1970). *J. Mol. Biol.* **53**, 159–162.
Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Miller, D. L., Hachmann, J. & Weissbach, H. (1971). *Arch. Biochem. Biophys.* **144**, 115–121.
Moldave, K. (1985). *Annu. Rev. Biochem.* **54**, 1109–1149.
Nagata, S., Tsunetsugu-Yokota, Y., Naito, A. & Kaziro, Y. (1983). *Proc. Nat. Acad. Sci., U.S.A.* **80**, 6192–6196.
Nagata, S., Nagashima, K., Tsunetsugu-Yokota, Y., Fujimura, K., Miyazaki, M. & Kaziro, Y. (1984). *EMBO J.* **3**, 1815–1830.
Olson, M. V., Loughney, K. & Hall, B. D. (1979). *J. Mol. Biol.* **132**, 387–410.
Perkins, D. (1949). *Genetics*, **34**, 607–626.
Qin, S., Moldave, K. & McLaughlin, C. S. (1987). *J. Biol. Chem.* **262**, 7802–7807.
Rose, M., Novick, P., Thomas, J., Botstein, D. & Fink, G. R. (1987). *Gene*, in the press.
Rotenberg, M. O. & Woolford, J. L. (1986). *Mol. Cell. Biol.* **6**, 674–687.
Rothstein, R. J. (1983). *Methods Enzymol.* **101**, 202–211.
Sacerdot, C., Dessen, P., Hershey, J. W. B., Plumbridge, J. A. & Grunberg-Manago, M. (1984). *Proc. Nat. Acad. Sci., U.S.A.* **81**, 7787–7791.
Saha, S. K. & Chakraburtty, K. (1986). *J. Biol. Chem.* **261**, 12599–12603.
Sanger, F., Nicklen, S. & Coulson, A. R. (1977). *Proc. Nat. Sci., U.S.A.* **74**, 5463–5467.
Sharp, P. M., Tuohy, T. M. F. & Mosurski, K. R. (1986). *Nucl. Acids Res.* **14**, 5125–5143.
Sherman, F., Fink, G. R. & Lawrence, C. W. (1971). *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sherman, F., Fink, G. R. & Hicks, J. (1982). *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Skogerson, L. & Engelhardt, D. (1977). *J. Biol. Chem.* **252**, 1471–1475.
Skogerson, L. & Wakatama, E. (1976). *Proc. Nat. Acad. Sci., U.S.A.* **73**, 73–76.
Song, J. M. & Liebman, S. W. (1987). *Genetics*, **115**, 451–460.
Southern, E. M. (1975). *J. Mol. Biol.* **98**, 503–517.
Surguchov, A. P., Smirnov, V. N., Ter-Avanesyan, M. D. & Inge-Vechtomov, S. G. (1984). *Physicochem. Biol. Rev.* **4**, 147–205.
Tapio, S. & Kurland, C. G. (1986). *Mol. Gen. Genet.* **205**, 186–188.
Teem, J. L., Abovich, N., Kaufer, N. F., Schwidinger, W. F., Warner, J. R., Levy, A., Woolford, J., Leer, R. J., van Raamsdonk-Duin, M. M. C., Mager, W. H., Planta, R. J. Schultz, L., Friesen, J. D., Fried, H. & Rosbash, M. (1984). *Nucl. Acids Res.* **12**, 8295–8312.
Thompson, R. C., Dix, D. B. & Karim, A. M. (1986). *J. Biol. Chem.* **261**, 4868–4874.
van Buul, C. P. J. J., Visser, W. & van Knippenberg, P. H. (1984). *FEBS Letters*, **177**, 119–124.
van Hemert, F. J., Amons, R., Pluijms, W. J. M., van Ormondt, H. & Moller, W. (1984). *EMBO J.* **3**, 1109–1113.
van Ness, B. G., Howard, J. B. & Bodley, J. W. (1978). *J. Biol. Chem.* **253**, 8687–8690.
van Noort, J. M., Kraal, I
Nyborg, J. & Clark, I
Sci., U.S.A. **81**, 3969
van Noort, J. M., Kraal, I
Acad. Sci., U.S.A. **82**
Vignais, M.-L., Woudt, L
W. H., Sentenac, A. &
6, 1451–1457.
Vijgenboom, E., Vink, T.
EMBO J. **4**, 1049–10
Wahl, G. M., Stern, M. &
Acad. Sci., U.S.A. **7**
Warner, J. (1982). In *Th*
Saccharomyces: *Me*
(Strathern, J., Jones
561, Cold Spring H
Spring Harbor, NY.

van Noort, J. M., Kraal, B., Bosch, L., la Cour, T. F. M., Nyborg, J. & Clark, B. F. C. (1984). *Proc. Nat. Acad. Sci., U.S.A.* **81**, 3969–3972.

van Noort, J. M., Kraal, B. & Bosch, L. (1985). *Proc. Nat. Acad. Sci., U.S.A.* **82**, 3212–3216.

Vignais, M.-L., Woudt, L. P., Wassenaar, G. M., Mager, W. H., Sentenac, A. & Planta, R. J. (1987). *EMBO J.* **6**, 1451–1457.

Vijgenboom, E., Vink, T., Kraal, B. & Bosch, L. (1985). *EMBO J.* **4**, 1049–1052.

Wahl, G. M., Stern, M. & Stark, G. R. (1979). *Proc. Nat. Acad. Sci., U.S.A.* **76**, 3683–3687.

Warner, J. (1982). In *The Molecular Biology of the Yeast* Saccharomyces: *Metabolism and Gene Expression* (Strathern, J., Jones, E. & Broach, J., eds), pp. 529–561, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Warner, J., Mitra, G., Schwindinger, W., Studeny, M. & Fried, H. (1985). *Mol. Cell. Biol.* **5**, 1512–1521.

Weiss, R. B. (1983). *Proc. Nat. Acad. Sci., U.S.A.* **81**, 5797–5801.

Weiss, R. B. & Gallant, J. (1983). *Nature (London)*, **302**, 389–393.

Weiss, R. B. & Gallant, J. (1986). *Genetics*, **112**, 727–739.

Weiss, R. B., Murphy, J. P. & Gallant, J. A. (1984). *J. Bacteriol.* **158**, 362–364.

Woudt, L. P., Smit, A. B., Mager, W. H. & Planta, R. J. (1986). *EMBO J.* **5**, 1037–1040.

Yamamoto, K. R. & Alberts, B. M. (1970). *Virology*, **40**, 734–744.

Zaret, K. S. & Sherman, F. (1982). *Cell*, **28**, 563–573.

Zengel, J., Archer, R. & Lindahl, L. (1984). *Nucl. Acids Res.* **12**, 2181–2192.

Edited by M. Gottesman

SYNAPSE 27:208–229 (1997)

Alzheimer Disease Hyperphosphorylated Tau Aggregates Hydrophobically

GEORGE C. RUBEN,[1]* THOMAS L. CIARDELLI,[2] INGE GRUNDKE-IQBAL,[3] AND KHALID IQBAL[3]
[1]*Department of Biological Sciences, Dartmouth College, Hanover, New Hampshire, 03755*
[2]*Department of Pharmacology and Toxicology, Dartmouth Medical School, Hanover, New Hampshire 03756*
[3]*New York State Institute for Basic Research in Developmental Disabilities, Staten Island, New York 10314*

KEY WORDS neurofibrillary tangles; MAP tau; circular dichroism; inverse temperature transition; transmission electron microscopy; freeze-dried vertically Pt-C replicated tau on silver filters

ABSTRACT The chemical interaction that condenses the hyperphosphorylated protein tau in Alzheimer's disease (AD P-tau) into neurofibrillary tangles and cripples synaptic transmission remains unknown. Only β-sheet, positive ion salt bridges between phosphates, and hydrophobic association can create tangles of just AD P-tau. We have correlated transmission electron microscope (TEM) images of tau aggregation with different percentages of β-sheet in aqueous suspensions of tau while using buffers that block dispositive or tripositive ionic bridges between intermolecular phosphates. Circular dichroism (CD) studies were performed at different temperatures from 5–85°C using AD P-tau, AD P-tau dephosphorylated with hydrofluoric acid (HF AD P-tau) or alkaline phosphatase (AP AD P-tau), and recombinant human tau with 3-repeats and two amino terminal inserts (R-39) and using bovine tau (B tau) isolated without heat or acid treatment. Secondary structure was estimated from CD spectra at 5°C using the Lincomb algorithm. Each preparation except one demonstrated an inverse temperature transition, T_t, in the CD at 197 nm. No correlation was found between β-sheet content and aggregation, leaving only hydrophobic interaction as the remaining possibility. Thirteen of 21 possible phophorylation sites in AD P-tau lie adjacent to positive residues in tau's primary structure. Occupation of five to nine phosphate sites on AD P-tau appears sufficient to reduce or neutralize tau's basic character. AD P-tau's hydrophobic character is indicated by its low inverse temperature transition, T_t. The T_t for AD P-tau was 24.5°C or 28°C, whereas for B tau with three phosphates it was 32°C, for unphosphorylated tau R-39 it was 38°C, and for dephosphorylated HF AD P-tau it was 37.5°C. The hydrophobic protein elastin and its analogs coalesce and precipitate at their T_t of 24–29°C, well below body temperature. We hypothesize that AD P-tau causes tangle accumulation by this mechanism. **Synapse 27:208–229, 1997.**

INTRODUCTION

During the progress of their disease, Alzheimer patients can develop a substantial presynaptic deficit. This problem can be partially overcome by treatment with acetylcholine esterase inhibitors (Francis et al., 1995). Beyond this deficit stage, synaptic communication between nerve cells collapses, followed eventually by nerve cell death. Alzheimer's disease (AD) has a complex etiology which probably includes genetic, environmental, and metabolic factors. Histologically, AD is characterized by the presence of intracellular neurofibrillary tangles of paired helical filaments (PHF) as well as extracellular β-amyloid in the brain parenchyma and in brain blood vessel walls (Kidd, 1964). Dementia in AD is associated with neurofibrillary degeneration; β-amyloid in the absence of neurofibrillary degeneration is not associated with clinical Alzheimer's disease (Barcikowska et al., 1989; Dickson et al., 1988; Katzman et al., 1988).

The microtubule associated protein (MAP) tau is recognized as the principal constituent in PHF and Alzheimer's neurofibrillary tangles (Goedert et al., 1992; Grundke-Iqbal et al., 1986a; Jakes et al., 1991). Tangles contain PHF, a triple-stranded 2.1 nm filament similar to tau polymer, and other amorphous structures contain-

Contract grant sponsor: NIH; Contract grant numbers: AG11054, NS18105, AG05892, AG08076, TW00507.

*Correspondence to: George C. Ruben, Dept. Biological Sciences, Dartmouth College, Hanover, NH 03755. E-mail: George.C.Ruben@Dartmouth.Edu

Received 23 July 1996; Accepted 30 December 1996

ing tau (Bancher et al., 1989; Köpke et al., 1993; Ruben et al., 1991, 1992, 1993, 1995). Tau is normally located in nerve cell axons associated with and stabilizing the microtubules required for axonal transport. Sequestration of tau into tangles and its removal from microtubules not only compromises axonal transport but cripples synaptic communication. Tau in PHF and in tangles has been shown to be abnormally hyperphosphorylated (Grundke-Iqbal et al., 1986b; Köpke et al., 1993). Abnormally phosphorylated Alzheimer's disease tau (AD P-tau) can be phosphorylated at 21 sites (Brion et al., 1991; Hasegawa et al., 1992; Iqbal and Grundke-Iqbal, 1995; Iqbal et al., 1989; Morishima-Kawashima et al., 1995a,b), whereas 12 sites can be phosphorylated on fetal tau (Watanabe et al., 1993; Morishima-Kawashima et al., 1995b), and normal adult tau can be phosphorylated only on five sites (Watanabe et al., 1993). Both adult and fetal tau stimulate microtubule assembly (Yoshida and Ihara, 1993), but AD P-tau cannot (Alonso et al., 1994, 1996; Iqbal et al., 1986). AD P-tau binds normal tau and not only prevents it from stimulating microtubule assembly (Alonso et al., 1994) but also depolymerizes microtubules by sequestering normal tau (Alonso et al., 1996).

The chemical interaction by which hyperphosphorylated tau associates in PHF and aggregates within tangles is not understood. For AD P-tau aggregation, only three interactions seem likely: 1) AD P-tau can associate with other tau by forming β-sheet, 2) dipositive or tripositive ionic bridges can occur between protein phosphate groups, or 3) hydrophobic associations can occur between AD P-tau. We have estimated the β-sheet content as well as other secondary structures in AD P-tau, dephosphorylated AD P-tau, recombinant human tau (R-39), and bovine tau (B tau) preparations using circular dichroism (CD). In the prosecution of these studies the dipositive ion chelating agents EDTA/EGTA or a phosphate buffer were always present to block divalent or trivalent ionic bridging. We reasoned that if the first and second interactions above were ruled out, then hydrophobic association (the third interaction) between AD P-tau would be implicated.

Using transmission electron microscopy (TEM), aggregation of the tau samples was observed in two ways. Samples were either negatively stained on thin carbon films or vertically replicated (Ruben, 1989) on 0.2 μm silver filters. The negative staining technique is excellent for identifying large tau aggregates but is unreliable in imaging small aggregates of a few tau monomers. To avoid this problem, we have used a new silver filter method which is able to observe small aggregates of tau. Tau apparently adhers to silver filters, and nonassociated tau can be removed by washing. It was in this fashion that AD P-tau aggregation was observed since its aggregates could not be removed by washing. All the tau preparations were freeze-dried and vertically replicated on silver filters using this new aggregation assay.

For many years the favored method of isolation of tau has used both an acid pH of ~2.7 and heating to 100°C, conditions that remove other microtubule associated proteins (MAPs) and result in a purified tau active in assembling tubulin dimers into microtubules (Grundke-Iqbal et al., 1986a; Lindwall and Cole, 1984a,b). Before the experiments reported here, we had generally assumed that tau in solution recovered any secondary structure needed to promote microtubule assembly or tau polymer formation (Ruben et al., 1991). However, formation of 2.1 nm tau polymers has consistently failed with tau treated with heat and acid, despite its ability to promote microtubule assembly. Our CD work indicates that taus isolated with acid and heat have secondary structures consisting mostly of unordered coil and variable amounts of β-sheet. We show that tau isolated without heat or acid treatment retains more of its predicted secondary structure (AD P-tau #1 and B tau), which we believe will be important for assembling triple-stranded 2.1 nm tau polymer filaments (Ruben et al., 1991) as well as forming PHF under physiological conditions.

METHODS AND MATERIALS

Tau isolation and handling for the circular dichroism spectra

Alzheimer's disease hyperphosphorylated tau (AD P-tau)

AD P-tau #1 was isolated from an AD brain by the procedure of Köpke et al. (1993), avoiding heat treatment, acid treatment, and lyophilization during its preparation. This sample was saved in 30% glycerol at −20°C over a weekend and dialyzed against 50 mM sodium phosphate. For CD experiments, this sample was filtered with a 0.45 μm Millipore (Bedford, MA) filter and then diluted 1:1 with 50 mM sodium phosphate (pH 6.5).

A second AD P-tau #2 was also isolated from an AD brain by the procedure of Köpke et al. (1993) by phosphocellulose chromatography and then precipitated with 80% ethanol to remove salts and lyophilized. This sample also avoided heat and acid treatment. This second sample was resuspended for CD in 25 mM sodium phosphate (pH 6.5) and filtered with a 0.45 μm Millipore centrifuge filter before its CD spectrum was taken. The filtered samples were used for preparing the freeze-etched and replicated samples on the silver filters. AD P-tau #1 was used for the CD spectra in Figures 1 and 2a and the inverse temperature transition in Figure 2e. AD P-tau #1 was also used in Figures 5A, 6A, and 8A,B. AD P-tau #2 imaged in Figures 9A–C, and its inverse temperature transition is displayed in Figure 2e. Both AD P-tau samples #1 and #2 appear in

Table I. The unfiltered AD P-tau #1 was also negatively stained in Figure 4.

Hydrofluoric acid dephosphorylated AD P-tau (HF AD P-tau)

AD P-tau #2 isolated as described above was methanol-precipitated prior to treatment with hydrofluoric acid. Two samples of 20–100 µg of this precipitated protein were combined with either 25 µl or with 50 µl of 48% hydrofluoric acid and incubated overnight at 4°C according to Greenberg et al. (1992). The samples were then lyophilized to remove hydrogen fluoride or hydrofluoric acid (HF). HF AD P-tau #1 and #2 were dissolved in 25 mM sodium phosphate (pH 6.5) and filtered with a 0.45 µm Millipore centrifuge filter before a CD spectrum was taken. The filtered HF AD P-tau #1 and #2 were used for preparing the freeze-etched and vertically replicated samples on the silver filters. HF AD P-tau #1 appears in Figure 5B, and HF AD P-tau #2 appears in Figures 6B and 10. HF AD P-tau #2 appears in the CD spectra in Figures 1 and 2b and the inverse temperature transition in Figure 2e. The secondary structures listed in Table I are the average of CD spectra of HF AD P-tau #1 and #2. Unfiltered HF AD P-tau was negatively stained, but no condensed tau was seen (figure not shown). Although R-39 tau was also reacted with HF under the same conditions, we did not continue with this experiment since we found out that HF AD P-tau was mostly unstructured coil, similar to R-39 tau before it was treated with HF. No change in the structure of the R-39 tau would have been seen as a result of treatment with HF.

Alkaline phosphatase treatment of AD P-tau (AP AD P-tau)

AD P-tau (200 µg/100 µl of 50 mM Tris, pH 8.6, 5 mM $MgCl_2$, 1 mM EGTA, and 5 mM PMSF) was dephosphorylated with alkaline phosphatase (500 units/ml for 3 h at 37°C). At the end of the incubation, the phosphatase was removed from the tau by boiling the reaction mixture for 5 min and sedimenting it at 15,000*g* for 20 min at 4°C. The supernatant was lyophilized, washed twice with 70% ethanol, and dried again in a vacuum concentrator. The material thus obtained was solubilized in 25 mM sodium phosphate, pH 6.5, filtered through a 0.45 µm mesh, and used for circular dichroism spectroscopy and for preparing samples on silver filters. Only AP AD P-tau's secondary structure is listed in Table I. The data for its inverse temperature transition was unsmoothable, so its T_t could not be obtained.

Recombinant human tau 39 (R-39 tau)

R-39 tau was isolated from *E. coli* as described previously (Singh et al., 1995), using both heat and acid treatments and subsequent lyophilization. Two samples, #1 and #2, were evaluated by CD after suspension in 25 mM sodium phosphate, pH 6.5, and filtration through a 0.45 µm Millipore centrifuge filter. The filtered samples were used to prepare the freeze-etched and replicated samples on the silver filters. R-39 tau filtered sample #2 was diluted 50% with trifluoroethanol for its CD spectrum. Filtered R-39 tau #2 was used in Figures 1, 2c,e, 7A, and 11. Unfiltered R-39 tau #1 was used in Fig. 5C. The deconvolution of R-39 tau #1 yielded the lowest error but was averaged with R-39 tau #2 which had similar secondary structure percentages.

Bovine tau (B tau)

Bovine brain, cleaned free of meninges, was homogenized in 9 vols of 20 mM Tris, pH 8.0, containing 0.32 M sucrose, 5 mM benzamidine, 10 mM B-mercaptoethanol, 5 mM EGTA, 0.5 mM $MgSO_4$, 5 mM B-glycerophosphate, 6 mM PMSF, 1 mM EDTA, 50 mM NaF, 1 mM sodium·vanadate, 0.1 mM chloroquinine, 10 nM soya bean trypsin inhibitor, 0.1 mg/ml tosyl arginine methionine, 5 µg/ml leupeptin, 2 µg/ml aportinine, and 1 µg/ml pepstatin. The homogenate was centrifuged at 100,000*g* for 30 min, and the supernatant obtained was precipitated at 30% and 50% ammonium sulfate saturation. The 30–50% ammonium sulfate fraction was subjected to gel filtration on a sephacryl-100 column equilibrated with 25 mM 2-[N-morpholino] ethane-sulfonic acid (MES), pH 6.8, containing 0.5 mM $MgCl_2$, 10 mM dithiothreitol, and 0.1 mM EDTA. The tau peak was pooled and subjected to phosphocellulose chromatography in the MES buffer. The column was eluted with a gradient of 0.0–0.5 M NaCl in the buffer. The tau peak was pooled, dialyzed against three changes of 50 mM MES, pH 6.9, and centrifuged at 80,000*g* for 30 min, and its supernatant was subjected to mono-S ion exchange FPLC (Pharmacia, Uppsala, Sweden) chromatography. Protein fractions eluted with 0.0–0.5 M NaCl were assayed for tau by Western blots. The fractions containing tau were pooled, concentrated, and buffer exchanged to 25 mM sodium phosphate, pH 6.5. This sample was filtered through a 0.45 µm filter before its CD spectrum was taken. The filtered sample was used for preparing the freeze-etched and replicated sample on a silver filter. The SDS-PAGE of this sample is shown as lane 2 in Figure 3. Tau is present along with other proteins below the tau bands. The presence of extra protein bands probably accounts for B tau's compromised inverse temperature transition above 35°C in Figure 2e and for its less than predicted α-helix and β-turn secondary structure based on its primary sequence (Table I). CD spectra of B tau appear in Figures 1 and 2d. TEM images of B tau appear in Figures 6C, 7B, and 12.

Circular dichroism

The CD spectra as a function of temperature were conducted as previously described in Ruben et al. (1991) using the method of Ciardelli et al. (1988) and an Instruments SA, Jobin Yvon Mark V circular dichro-

graph (Longjumeau, France). All of the samples were filtered through a 0.45 μm Millipore centrifuge filter before their spectra were recorded so that the 185 nm signal could be observed and the 195–245 nm signal could be smoothed. The concentrations of the filtered protein solutions were measured using the methods of Bensadoun and Weinstein (1976). Each spectrum was normalized to the moles of protein contained and its theta spectrum computed. Theta values from 195–245 nm at 1 nm intervals were employed to estimate percent aromatics and protein secondary structures of α-helix, β-turn, unstructured coil, and β-sheet using the Lincomb algorithm (Perczel et al., 1991). Since percentages in Table I were based only on protein secondary structure, the aromatics were removed and the secondary structure percentages recalculated. Since all the tau secondary structures were calculated similarly, we felt that even if systematic error were present altering absolute secondary structure percentages, differences would be real. Two samples or more of each type of tau were deconvoluted, and the percentages of each secondary structure were averaged. The uncertainty of each secondary structure percentage was the standard error of the mean. The inverse temperature transition was calculated at 197 nm because it is representative of the coil maxima at 196 nm and is more readily smoothed to remove noise than 196 nm.

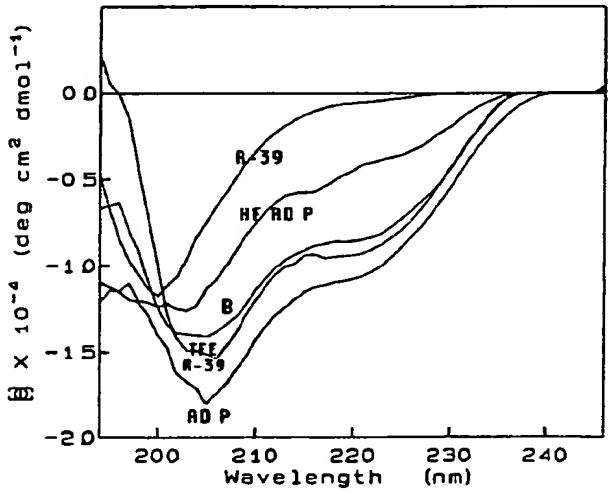


Fig. 1. CD spectra of various tau at 5°C in sodium phosphate buffer, pH 6.5. Using the Lincomb algorithm, we deconvoluted these CD spectra into components of α-helix, β-turn, unstructured coil, β-sheet, and aromatics. In Table I the aromatics were removed, and the percentage of each secondary structure was recomputed. CD spectra of AD P-tau #1 (AD P), human R-39 tau #2 in 50% trifluoroethanol (TFE R-39), bovine tau isolated without heating or acid treatment (B), AD P-tau #2 dephosphorylated with hydroflouric acid (HF AD P), and R-39 human tau #2 (R-39) were plotted. Alkaline phoshatase treated AD P-tau was not included.

Preparations of transmission electron microscopy specimens

Negative staining preparations

Unfiltered AD P-tau #1 was negatively stained with 2% uranyl acetate (pH 3.8) on ~10 nm double indirect

TABLE I. Circular dichroism estimate of β-sheet secondary structure of the microtubule associated protein tau correlated with degree of aggregation by transmission electron microscopy

Samples	α-helix (%)	β-turn (%)	Coil (%)	β-sheet (%)	Aggregation TEM[8]
AD P-tau[1] (sample #1)	0.0	36.8 ± 0.7	57.1 ± 0.7	6.8 ± 0.9	++++++
AD P-tau[1] (sample #2)	0.0	10.4 ± 1.3	81.7 ± 1.2	7.8 ± 0.1	++++++
AP AD P-tau[2]	0.0	0.0	58.2 ± 0.6	41.8 ± 0.6	≤++
HF AD P-tau[3] (sample #1 and #2)	3.1 ± 1.1	0.0	68.0 ± 3.5	28.9 ± 2.5	≤ + 1/2
R-39 tau[4] (sample #1 and #2)	0.0	0.0	80.3 ± 1.6	19.7 ± 1.6	≤+
TFE R-39 tau[5]	15.6 ± 0.3	18.2 ± 0.3	66.2 ± 0.3	0.0	0
B tau[6]	5.1 ± 0.8	23.1 ± 0.6	64.3 ± 0.2	7.4 ± 0.7	0
Human tau, sequence prediction[7]	20.0	38.0	42.0	0.0	
Bovine tau, sequence prediction[7]	23.0	38.0	39.0	0.0	
Bovine elastin, sequence prediction[7]	21.0	35.0	44.0	0.0	

[1]Alzheimer's disease hyperphosphorylated tau (AD P-tau) isolated in the absence of heating or acid pH by the methods of Köpke et al. (1993). The table values are the average of similar secondary structure values derived from CD spectra at 5, 15, and 35°C for sample #1. Sample #2, which had been precipitated with 80% ethanol, appears to have increased coil and decreased β-turn content. Aggregation was equivalent to sample #1.
[2]Alkaline phosphatase treatment of AD-P tau removes ~90% of its phosphate groups (unpublished observation).
[3]Alzheimer's disease hyperphosphorylated human tau dephosphorylated with hydroflouric acid treatment (HF AD P-tau) at 4°C according to Greenberg et al. (1992). This CD secondary structure percentage is the average of sample #1 and #2 secondary structures at 5°C.
[4]Recombinant human tau 39, R-39, isolated as described previously (Singh et al., 1995) using heat and acid pH. The average secondary structures of samples #1 and #2. The CD spectra were taken at 5°C.
[5]The recombinant human tau 39, R-39, solution from sample #2 (see footnote 4) was diluted with trifloUoroethanol (TFE) until the solution was 50% TFE.
[6]Bovine tau was isolated without using heat, acid treatment, or lyophilization.
[7]See Ruben et al. (1991) for the derivation of these estimates.
[8]Aggregation scale ranges from most aggregated (+ + + + +) to least aggregated (0). This information was derived from TEM images of tau aggregation on silver filters.

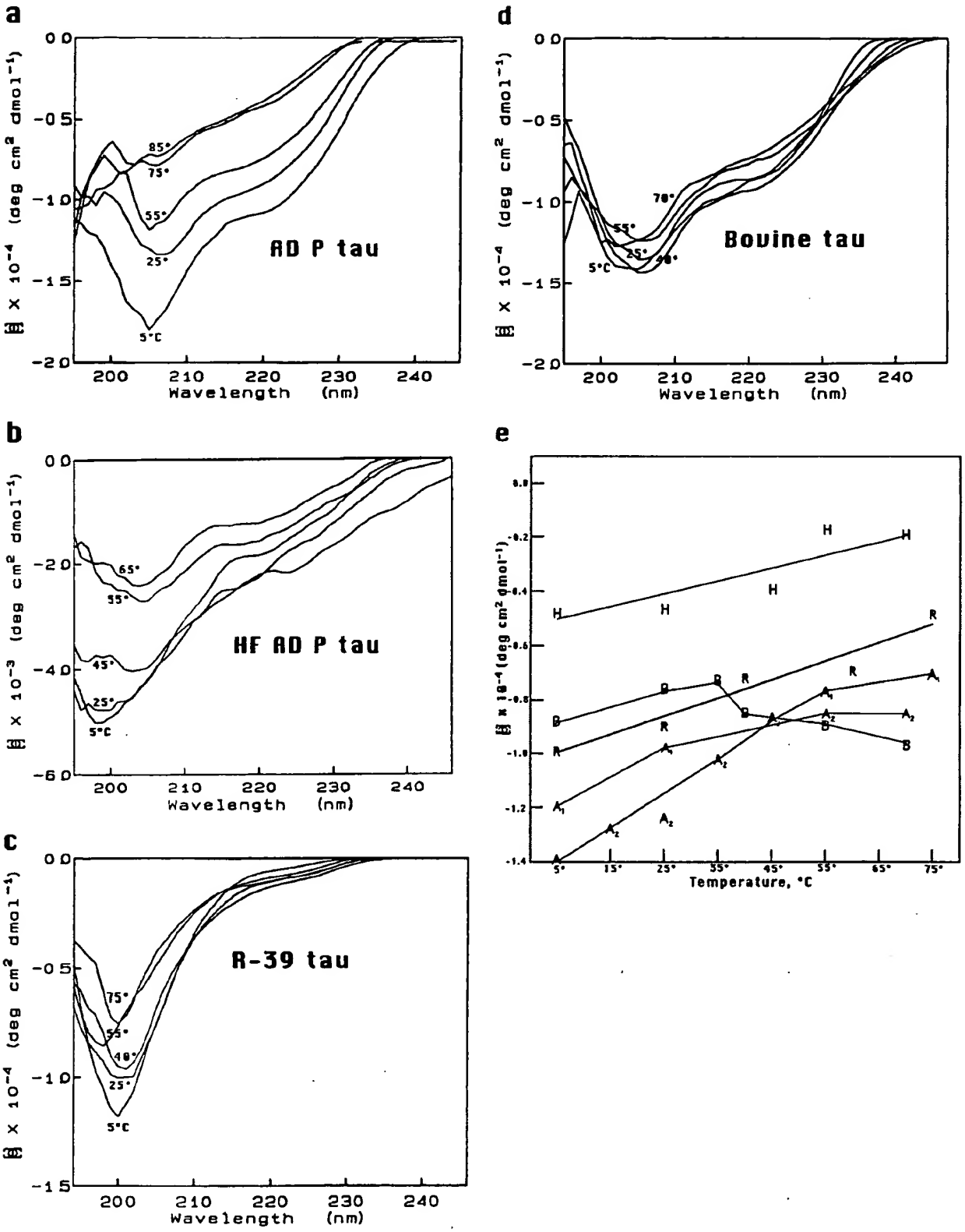


Fig. 2 (Legend on facing page.)

carbon film treated with 25% glutaraldehyde vapor at room temperature for 3 h (Ruben et al., 1988) and appears in Figure 4. Unfiltered HF AD P-tau and AP AD P-tau were also prepared, but no aggregation was apparent. Only HF AD P-tau's freeze-etch results are presented since this sample was assumed to be completely dephosphorylated (Greenberg et al., 1992).

Sample preparation on 0.2 μm porosity silver filters for freeze-drying and vertical Pt-C replication

Tau solutions (35–250 μg/ml) in 50 μl aliquots were placed at the center of the 13 mm 0.2 μm porosity silver disc at room temperature (~20°C) on the spin axis of a table top centrifuge and spun until the drop spread out radially from the center of the filter and thinned as it disappeared over the edge of the silver filter. The filter was removed from the centrifuge and washed thoroughly with distilled water in a Boyden chamber filter unit (Ruben et al., 1992). The filter was plunge-frozen in liquid propane cooled with liquid nitrogen and later freeze-dried at −80°C for 1.5 h in a modified Balzers 300, replicated with 0.93–1.0 nm Pt/C and coated with 12.1–12.8 nm rotary deposited carbon applied in two steps (Ruben, 1989). The replica was removed from the filter by floating the silver filter on a saturated potassium cyanide solution overnight or until the silver was dissolved.

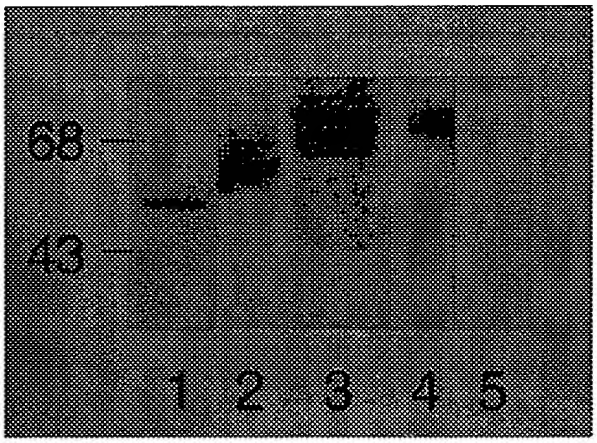


Fig. 3. Sodium dodecyl sulfate–polyacrylamide gel electrophoretic pattern of R-39 tau, B tau, and AD P-tau. Lane 1: Recombinant human tau R-39. Lane 2: Bovine tau. Lanes 3–5: Alzheimer disease abnormally phosphorylated tau (AD P-tau). Lanes 1–3: Comassie blue stained pattern. Lanes 4,5: Western blot developed with phosphorylation-dependent monoclonal antibody Tau-1; lanes 4 and 5 treated and untreated, respectively, with alkaline phosphatase on the blot prior to immunostaining.

This reaction is driven by air oxidation and works even in a covered petri dish (Ruben et al., 1996). The replica was floated on a saturated KOH solution overnight to remove the sample and then washed on distilled water for 24 h before being picked up on bare 300 mesh copper grids. A silver filter without any sample was also replicated in the Balzers 300 with 0.93 nm Pt-C and rotary carbon coated in two steps with 11.6 nm (Ruben, 1989). The silver filter was digested from beneath the replica as described and mounted on a bare 300 mesh copper grid.

Transmission electron microscopy and photographic methods

The replicas and negatively stained samples were examined with a JEM 100cx with a lanthanum hexaboride filament, using a 400 μm condenser and a 40 μm, 60 μm, or a 120 μm objective apertures. Both 80 kv and 100 kv accelerating voltages were used. The replicas (0.98–1.0 nm Pt-C) and negatively stained sample images were not better than 0.6–0.7 nm resolution. The 0.93 nm Pt-C film on the bare silver filter was not better than ~0.45 nm resolution. The negatively stained images were printed directly from the TEM negatives. The TEM negatives from the replicas were used to make reversal negatives as described before (Ruben, 1989). The reversal negatives were printed on 8″ × 10″ Ilford Multigrade II fiber based paper enlarged 2.5 times on a Durst 1200 point source enlarger with an

Fig. 2. **a:** CD spectrum of AD P-tau at 5, 25, 55, 75, and 85°C. Circular dichroism spectrum of ~160 μg/ml of AD P-tau #1 in 50 mM sodium phosphate, pH 6.5, as a function of temperature from 5–85°C. The mean residue ellipticity (Θ) was based on the human tau sequence and an average amino acid molecular weight of 104. This AD P-tau has a mixed phosphorylation state of ~7.57 ± 1.54 mol P/396.5 amino acid tau (average tau sequence of six alternately spliced tau isoforms with a mol. wt. of 41,318 g). **b:** CD spectrum of HF AD P-tau at 5, 25, 45, 55, and 65°C. Circular dichroism spectrum of ~50 μg/ml of HF AD P-tau, sample #2, in 25 mM sodium phosphate, pH 6.5, as a function of temperature from 5–65°C. The mean residue ellipticity (Θ) was based on the human tau sequence and an average amino acid molecular weight of 104. This HF AD P-tau has almost all of the phosphates removed with <1 mol P/396.5 amino acid tau. **c:** CD spectrum of human R-39 tau at 5, 25, 40, 55, and 75°C. Circular dichroism spectrum of ~250 μg/ml of R-39 tau (sample #2) in 25 mM sodium phosphate, pH 6.5, as a function of temperature from 5–75°C. The mean residue ellipticity (Θ) was based on the human tau sequence and an average amino acid molecular weight of 103.9. R-39 tau is unphosphorylated (410 amino acids, mol wt. 42,603 g/mol). **d:** CD spectrum of bovine tau (B tau) at 5, 25, 40, 55, and 70°C. Circular dichroism spectrum of ~75 μg/ml of bovine tau in 25 mM sodium phosphate, pH 6.5, as a function of temperature from 5–70°C. The mean residue ellipticity (Θ) was based on the bovine tau sequence and an average amino acid molecular weight of 103.4. This B tau has a mixed phosphorylation state averaging ~3.1 ± 0.4 mol P/384 amino acid tau (average tau sequence of six alternately spliced tau isoforms with a mol. wt. of 39,706 g). This sample was isolated without heat, acid, or lyophilization treatment to preserve its native secondary structure. **e:** Inverse temperature transition, T_t, in AD P-tau (A), HF AD P-tau (H), human R-39 tau (R), and bovine tau (B). The theta values at 197 nm as a function of temperature from 5–75°C. The inverse temperature transition in bovine tau has been previously described by Ruben et al. (1991). Dephosphorylated AD P-tau, HF AD P-tau (H), and R-39 tau (R) show inverse temperature transitions. AD P-tau #1 (A_1) and AD P-tau #2 (A_2) have inverse temperature transitions which are complete at about 55°C. The bovine tau (B) isolated without being subjected to heating, acid pH, or lyophilization has an inverse temperature transition from 5–35°C, but above this temperature theta at 197 nm decreases, indicating the formation of coil with increased temperature.

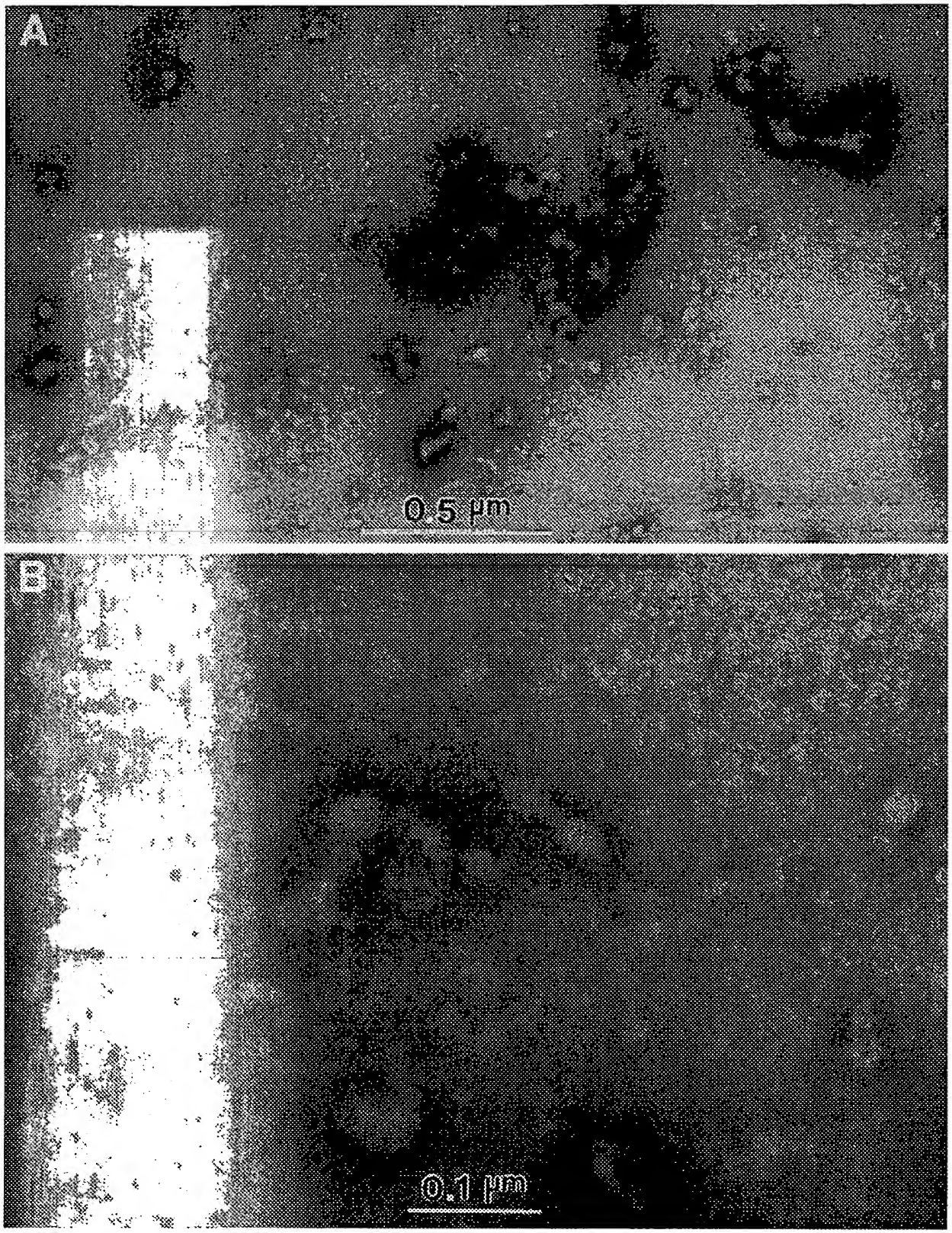


Fig. 4. Negatively stained AD P-tau #1 (160 μg/ml) in 50 mM $NaHPO_4$ (pH 6.5) negatively stained with 2% uranyl acetate (pH 3.8) on ~10 nm double indirect carbon films. **A:** AD P-tau forms globular condensates as well as extended condensates. The globular particulate structures range from ~7 nm to ~84 nm. The less regular condensates are larger, and they are as long as ~600 nm and as wide as ~200 nm. The surface of the carbon film also appears to be covered with thinner, less prominent condensed structures. No paired helical filaments were found in this preparation. ×70,000. **B:** A large L-shaped irregular structure is ~221 nm long with widths of 25–42 nm. The irregular long structure to its right is ~235 nm long and varies in width from 12.6–29.4 nm and contains three globular sections. The carbon film surface is covered with thin, irregularly shaped condensates as in A. ×245,000.

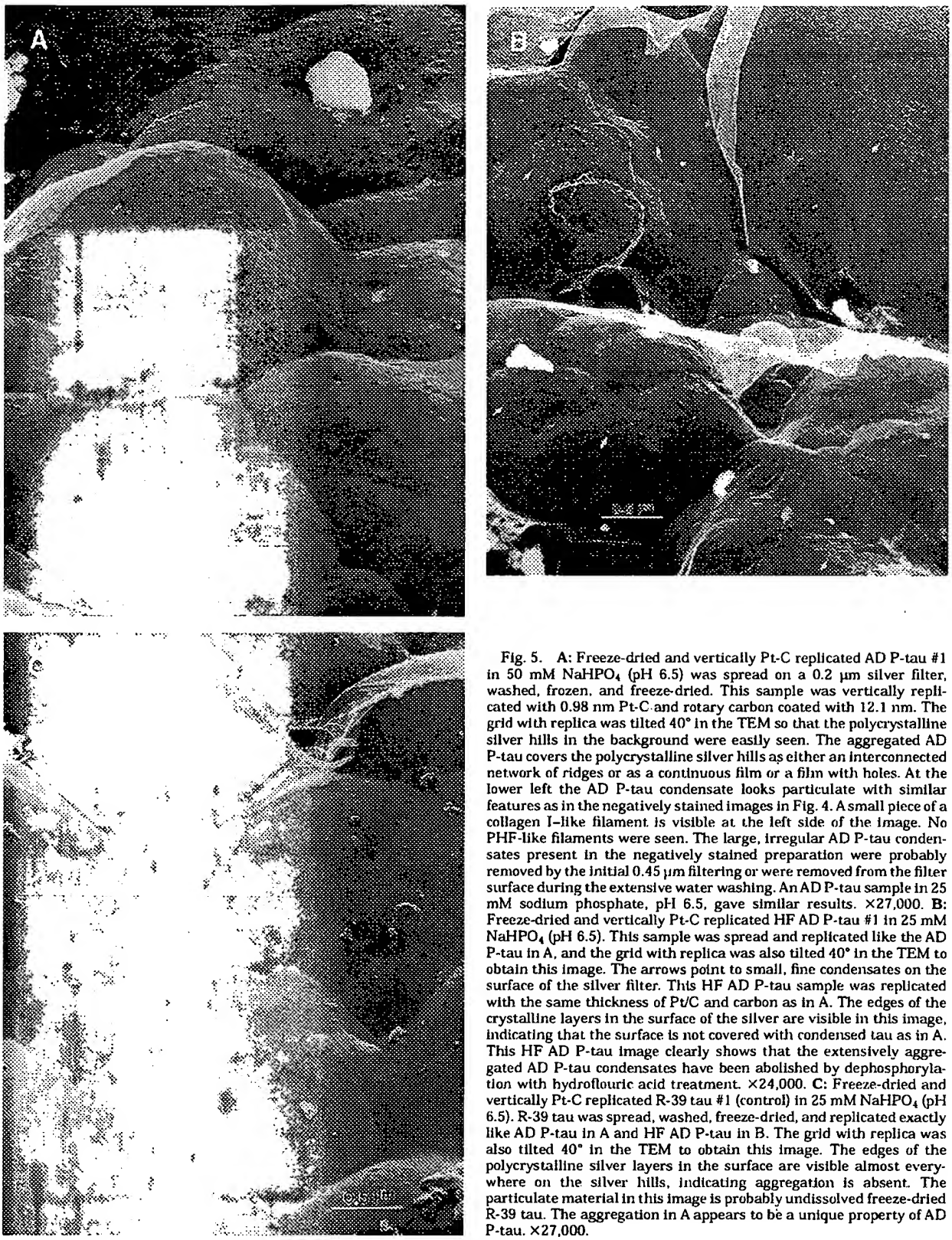


Fig. 5. **A:** Freeze-dried and vertically Pt-C replicated AD P-tau #1 in 50 mM $NaHPO_4$ (pH 6.5) was spread on a 0.2 μm silver filter, washed, frozen, and freeze-dried. This sample was vertically replicated with 0.98 nm Pt-C and rotary carbon coated with 12.1 nm. The grid with replica was tilted 40° in the TEM so that the polycrystalline silver hills in the background were easily seen. The aggregated AD P-tau covers the polycrystalline silver hills as either an interconnected network of ridges or as a continuous film or a film with holes. At the lower left the AD P-tau condensate looks particulate with similar features as in the negatively stained images in Fig. 4. A small piece of a collagen I-like filament is visible at the left side of the image. No PHF-like filaments were seen. The large, irregular AD P-tau condensates present in the negatively stained preparation were probably removed by the initial 0.45 μm filtering or were removed from the filter surface during the extensive water washing. An AD P-tau sample in 25 mM sodium phosphate, pH 6.5, gave similar results. ×27,000. **B:** Freeze-dried and vertically Pt-C replicated HF AD P-tau #1 in 25 mM $NaHPO_4$ (pH 6.5). This sample was spread and replicated like the AD P-tau in A, and the grid with replica was also tilted 40° in the TEM to obtain this image. The arrows point to small, fine condensates on the surface of the silver filter. This HF AD P-tau sample was replicated with the same thickness of Pt/C and carbon as in A. The edges of the crystalline layers in the surface of the silver are visible in this image, indicating that the surface is not covered with condensed tau as in A. This HF AD P-tau image clearly shows that the extensively aggregated AD P-tau condensates have been abolished by dephosphorylation with hydrofluoric acid treatment. ×24,000. **C:** Freeze-dried and vertically Pt-C replicated R-39 tau #1 (control) in 25 mM $NaHPO_4$ (pH 6.5). R-39 tau was spread, washed, freeze-dried, and replicated exactly like AD P-tau in A and HF AD P-tau in B. The grid with replica was also tilted 40° in the TEM to obtain this image. The edges of the polycrystalline silver layers in the surface are visible almost everywhere on the silver hills, indicating aggregation is absent. The particulate material in this image is probably undissolved freeze-dried R-39 tau. The aggregation in A appears to be a unique property of AD P-tau. ×27,000.

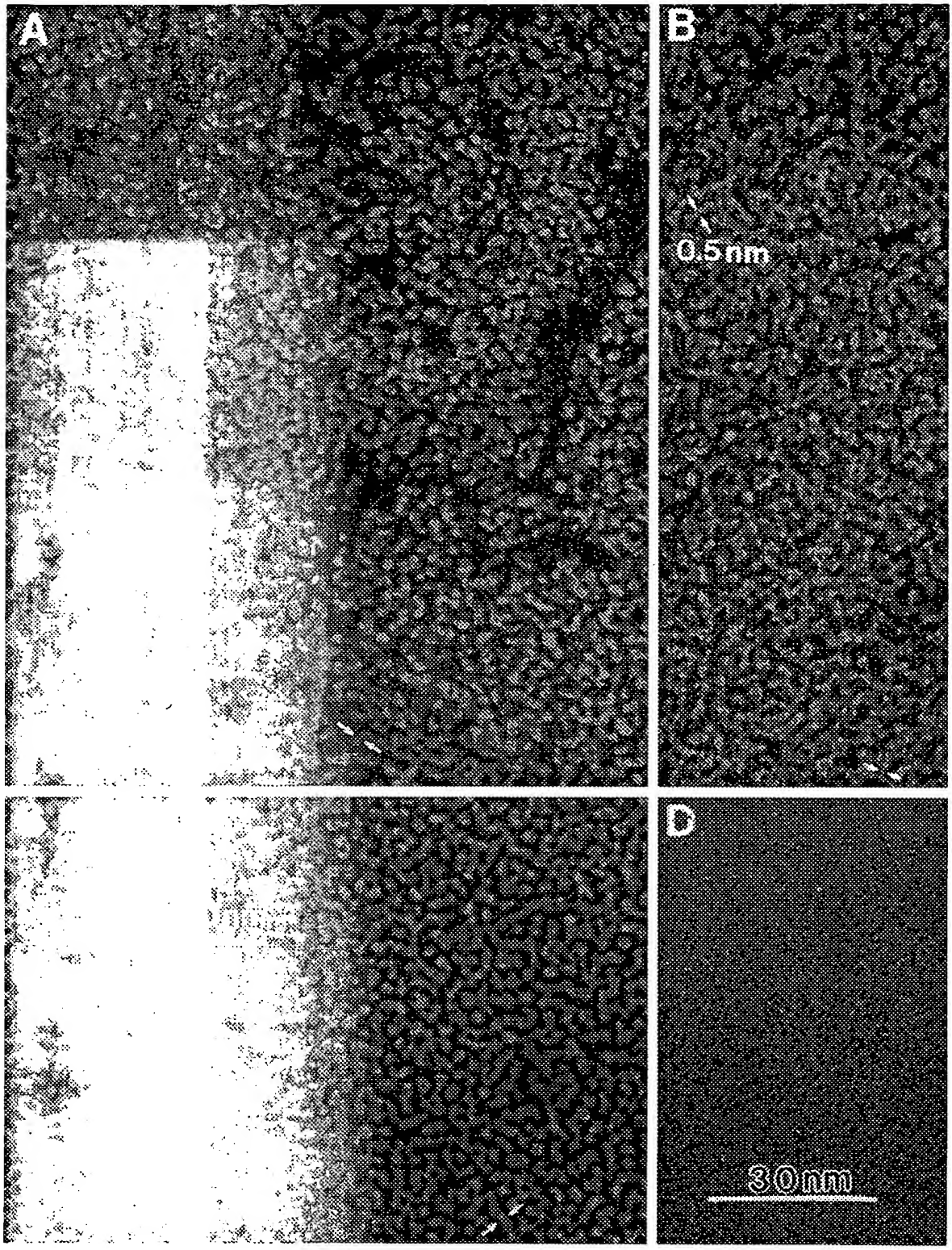


Fig. 6 (Legend on facing page.)

apochromatic 150 mm lens. The 1 million times enlargements were made with an apochromatic 80 mm lens.

RESULTS

Circular dichroism of tau

The secondary structures of tau preparations reported here were strongly influenced by their method of preparation. The most widely used purification method subjects tau to pH 2.7 and a brief exposure to 100°C (Grundke-Iqbal et al., 1986a; Lindwall and Cole, 1984a,b) which not only preserves its ability to promote microtubule assembly but produces a tau that is mostly unstructured coil. Hydrogen fluoride treatment of AD P-tau (HF AD P-tau) removes close to 100% of the phosphate groups, but it also removes internal hydrogen bonds producing mostly random coil. Lyophilization of samples containing mostly coil produces a variable but substantial β-sheet content. Tau preparations that more closely represent the native secondary structure were isolated in the absence of acid pH and heating. AD P-tau #1 and bovine tau (B tau) were preparations of this type. AD P-tau #2, unlike AD P-tau #1, was treated with alcohol before being lyophilized, which probably accounts for the reduction in its β-turn content.

CD spectra of R-39 tau alone and in 50% trifluoroethanol, AD P tau #1, HF AD P tau, and a nonconventionally isolated bovine tau (B tau) are shown in Figure 1. The percentages of α-helix, β-turns, coil, and β-sheet secondary structure for each of these samples are reported in Table I along with AD P-tau #2 and an alkaline phosphatase dephosphorylation of this same AD P tau #2 (AP AD P-tau) not shown in Figure 1. AD P-tau #1 with the highest level of aggregation had 36.8% β-turn, 6.8% β-sheet, and the rest of its secondary structure in the form of unordered coil. AD P-tau #2 was isolated similarly to AD P-tau #1 except that it was precipitated with 80% ethanol prior to lyophilization. AD P-tau #2 had a β-turn content of 10.4% with β-sheet remaining unchanged at 7.8%, and the unordered coil correspondingly increased. This sample was also highly aggregated. Dephosphorylation of AD P-tau using alkaline phosphate treatment removed ~90% of the AD P-tau phosphates (unpublished observation); AP AD P-tau was boiled and alkaline phosphatate–precipitated, and the supernatant was lyophilized and washed twice with ethanol. A second method utilized HF to remove all the phosphates (Greenberg et al., 1992): AD P-tau was methanol-precipitated and treated at 4°C with 48% HF overnight and then lyophilized. As shown in Table I, HF AD P-tau contained 3.1% α-helix, no β-turns, and 28.9% β-sheet, with the rest unordered coil. The AP AD P-tau contained 41.8% β-sheet and no β-turns. The AP AD P-tau was only slightly more aggregated than the HF AD P-tau and is not shown. Both AP AD P-tau and HF AD P-tau preparations were markedly less aggregated than the AD P-tau preparations. R-39 tau isolated with acid and heat and then lyophilized contained 19.7% β-sheet, with the rest unordered coil, but, when combined with 50% trifluoroethanol to stabilize α-helix, its secondary structure was 15.6% α-helix, 18.2% β-turn, with no β-sheet. R-39 tau in 50% trifluoroethanol was unaggregated but was lightly aggregated in the absence of trifluoroethanol. Unaggregated bovine tau (B tau) was isolated without heat or acid treatment, contained 5.1% α-helix, 23.1% β-turns, and 7.4% β-sheet, with the rest unordered coil; we had expected higher values for α-helix and β-turns which were likely reduced by the presence of protein contaminants (lane 2, Fig. 3).

Figure 2a–d shows the change in CD spectra accompanying an increase in temperature from 5°C to 65–85°C for AD P-tau #1, HF AD P-tau #2, R-39 tau, and B tau. Each family of curves has a theta value at 197 nm recorded in Figure 2e as a function of temperature. Half the total change in theta with temperature at 197 nm is the inverse temperature transition (T_i). The change in AD P-tau #2 (A_2) theta with temperature was also included in Figure 2e. Both AD P tau #1 (A_1) and #2 (A_2) show a T_i at 20°C (50 mM sodium phosphate, pH 6.5)

Fig. 6. A: High magnification image of freeze-dried, vertically replicated AD P-tau #1 in 50 mM $NaHPO_4$ (pH 6.5) spread on a 0.2 μm silver filter, washed, frozen, and freeze-dried. This sample was vertically replicated with 0.93 nm Pt-C and rotary carbon coated with 12.1 nm. This image shows condensed AD P-tau forming ridges 7–20 nm wide that stand above the surface with the ends of the AD P-tau monomers protruding from the ridges. The width of the monomers, ~1 nm, was corrected for the metal coating by subtracting 0.45 nm from the metal film thickness and subtracting 0.53 nm of Pt-C from the width of each metal-coated monomer. There are also AD P-tau lying horizontally on the surface with the silver surface visible between molecules adjacent to the ridges. The ridges show none of the periodicity of PHF. The ridges also form branching junctions with three or four (Fig. 5A) connecting ridges, unlike normal PHF which do not branch. ×1,000,000. B: High magnification image of freeze-dried, vertically replicated HF AD P-tau #2 (50 μg/ml) in 25 mM $NaHPO_4$ (pH 6.5) spread on a 0.2 μm silver filter, washed, frozen, and freeze-dried. This sample was vertically replicated with 1.0 nm Pt-C and rotary carbon coated with 12.8 nm. By dephosphorylating with HF, we abolished the AD P ridges in A. The HF AD P-tau shown here contains individual chains that are ~0.5 nm wide after correcting for the metal coating (0.55 nm of Pt-C was subtracted from the Pt-C coated filaments). The tau in this image are irregular, with a width of ~0.5 nm, the approximate diameter of an extended amino acid chain. Many of the tau are associated since the small aggregates frequently branch and separate into ~0.5 nm chains. Some tau molecules double back on themselves, suggesting they are self-associated. ×1,000,000. C: High magnification image of freeze-dried, vertically replicated Bovine tau (~75 μg/ml) in 25 mM $NaHPO_4$ (pH 6.5) spread on a 0.2 μm silver filter, washed, frozen, and freeze-dried. This sample was vertically replicated with 0.93 nm Pt-C and rotary carbon coated with 12.1 nm. The surface of the silver filter is coated with short sections of bovine tau which visualized in stereo images extend above the surface towards the observer. Where the tau extend parrallel with the surface, their width is often ~1.0 nm after correcting for the Pt-C coating (0.53 nm is subtracted from the Pt-C coated chains). These tau appear unassociated with other tau. ×1,000,000. D: Control silver filter surface coated with 0.93 nm Pt-C and 11.6 nm rotary carbon. The vertical Pt-C deposition on silver produces a metal-coated surface with very short metal chains of ~0.5 nm in width. The width of these metal film structures is less than half to a third the width of any of the Pt-C coated tau monomers in A–C. ×1,000,000.

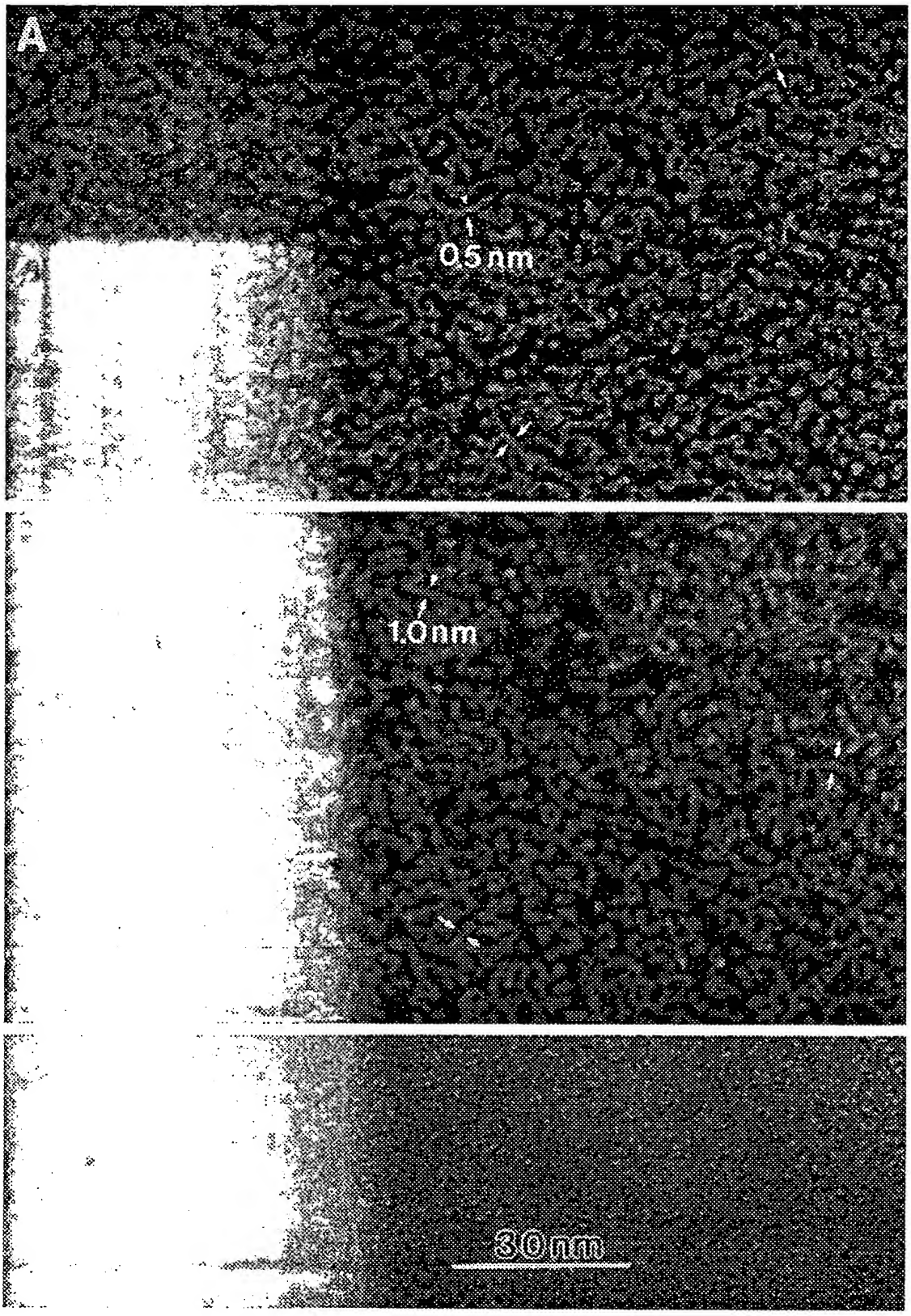


Fig. 7 (Legend on facing page.)

and 28°C (25 mM sodium phosphate, pH 6.5), respectively, where the 50 mM phosphate concentration is estimated to have reduced T_t by 4–5°C (Urry, 1993). HF AD P-tau (H, T_i = 38°C) and R-39 tau (R, T_t = 37.5°C) in Figure 2e both continue to increase beyond 65–70°C just as the original bovine tau (T_t = 32°C) (Ruben et al., 1991) isolated like R-39 tau. The only tau with a truncated (5–35°C) inverse temperature transition is bovine tau (B), where the decrease in theta above 35°C probably results from the presence of protein impurities (lane 2, Fig. 3) rather than being a property of B tau itself. It should be noted that tau samples undergoing the inverse temperature transition which contain more unstructured coil also show a greater change in the coil signal at 197 nm.

Transmission electron microscopy of Alzheimer's disease abnormally phosphorylated tau (AD P-tau) and dephosphorylated HF AD P-tau

High resolution TEM was primarily used to estimate the aggregation of tau after spreading each sample on a silver filter, followed by thorough washing to remove unaggregated tau with room-temperature water. Negative staining revealed aggregation only in AD P-tau samples. Estimates of total aggregation based on vertically replicated tau images are included in Table I. The characteristics of sample aggregation are described below. The aggregation of AD P-tau, low in β-sheet, is more extensive and clearly different from aggregation in tau samples high in β-sheet. Tau samples with little or no β-turns but high in unordered coil had extended chain diameters of ~0.5 nm. AD P-tau #1 and B tau which contained substantial β-turns (36.8% and 23.1%) and coil had diameters of ~1.0 nm, the cylindrical diameter of β-spiral since there was either none or very little α-helix present (Ruben et al., 1991). The only exception to this pattern was the R-39 tau in 50% trifluoroethanol (TFE). This sample contained 18.2% β-turn, 15.6% α-helix, with the rest coil, but its chain diameter was mixed and alternated between short regions of ~1.0 nm diameter and longer regions of ~0.5 nm chain diameter. It appears that 50% TFE does not stabilize a β-spiral structure like it does α-helix.

The aggregation of AD P tau #1 and #2 in low and high magnification images as well as stereo images has been well documented. Negatively stained unfiltered AD P-tau #1 is shown in Figure 4A,B. This specimen has a range of irregular complex structures as large as 200 nm × 600 nm which do not resemble paired helical filaments (PHF), and no PHF have been observed in this preparation. Higher magnification (Fig. 4B) shows not only large structures (25–42 nm × ~221 nm) but contains many more subtle, irregular film- or sheet-like structures and particles. Vertically replicated AD P-tau #1 is shown in Figure 5A. It was prefiltered with a 0.45 μm Millipore centrifuge filter for the CD spectra and later spread on a 0.2 μm silver filter with polycrystalline silver hills which form the background beneath tau and its aggregation structures. The sample and filter were thoroughly washed with distilled water. Only aggregated AD P-tau adhering to the filter surface remained. Figure 5A lacks the larger agglomerations seen by negative staining, but it nevertheless clearly shows a network of condensed tau with junctions of three to four connections. These aggregated networks are also connected to continuous films of AD P-tau. This aggregation was likely present in the original solution of 25 mM or 50 mM sodium phosphate, pH 6.5, and remained after washing; this sample was never lyophilized before it was spread on the silver filter. Figure 5B shows lyophilized HF AD P-tau that had been resuspended in 25 mM sodium phosphate, pH 6.5, and filtered with an 0.45 μm Millipore centrifuge filter before the CD spectrum was recorded and this sample prepared. Aliquots of 50 μl were later spread and washed on 0.2 μm silver filters. Prior to filtration HF AD P-tau was also negatively stained like AD P tau, but no aggregation was apparent (not shown). The polycrystalline planes in the filter surface are easily seen in Figure 5B, and they are unobscured in contrast to the silver filter in Figure 5A. Only small, isolated branched figures of aggregated HF AD P-tau remain, and the extensive AD P-tau aggregation in Figure 5A was abolished. Alkaline phosphatase–dephosphorylated AD P-tau also removed aggregation (not shown). Figure 5C shows the recombinant R-39 tau, lyophilized, resuspended in 25 mM sodium phosphate, pH 6.5, and filtered for the CD spectra and for TEM. Large pieces of undissolved R-39 tau are present, but the polycrystaline planes in the silver filter are easily observed and are not obscured by aggregated tau. Undissolved R-39 tau does not fall under our definition of aggregation and

Fig. 7. A: High magnification image of freeze-dried, vertically replicated R-39 tau #2 (~250 μg/ml) in 25 mM $NaHPO_4$ (pH 6.5) spread on a 0.2 μm silver filter, washed, and frozen for freeze-drying. This sample was vertically replicated with 1.0 nm Pt-C and rotary carbon coated with 12.1 nm. The R-39 tau is routinely isolated using both 100°C for ~5 min and an acid pH of 2.7. Where single fine filaments are visible (arrows), they are frequently ~0.5 nm wide (0.55 nm is subtracted from the filaments to remove the Pt-C coating). There are also fine filaments which appear to be self-associated and associated with other tau molecules. ×1,033,000. B: High magnification image of freeze-dried, vertically replicated Bovine tau (~75 μg/ml) in 25 mM $NaHPO_4$ (pH 6.5) spread on a 0.2 μm silver filter, washed, and frozen for freeze drying. This sample was vertically replicated with 0.98 nm Pt-C and rotary carbon coated with 12.1 nm. In contrast to R-39 tau in A, bovine tau was isolated without using heat, acid pH, or lyophilization and contains α-helix and β-turns as part of its secondary structure (see Table I). The bovine tau monomer fine filaments widths are larger (~1.0 nm) than the chain widths in the R-39 image (~0.5 nm) (0.53 nm is subtracted from the Pt-C coated chains). ×1,033,000. C: Control silver filter surface coated with 0.93 nm Pt-C and 11.6 nm rotary carbon. The vertical Pt-C deposition on silver produces a metal-coated surface with very short metal chains of ~0.5 nm in width. The width of these metal film structures is less than half to a third the width of any of the Pt-C coated tau monomers in A or B. ×1,033,000.

was therefore discounted. Figure 7A shows higher magnification images of R-39 tau.

At high magnification, Figure 6A shows a small area of the AD P-tau #1 network previously shown at lower magnification in Figure 5A. Tau molecules adhere to the surface between the network of ridges. The AD P-tau in the network appears associated with adjacent AD P-tau, with their long axes oriented approximately parallel; as a consequence, the ~1.0 nm diameter AD P-tau monomers protrude from the network ridges. Unlike PHF, the network ridges are of variable thickness, are not helical, and form junctions with three and four connecting ridges. The hydrofluoric acid–dephosphorylated AD P-tau in Figure 6B (HF AD P-tau) shows tau on the surface of the silver filter with the extensive aggregation of AD P-tau absent. The tau monomer chains average ~0.5 nm in diameter, the same as that of random coil. The HF AD P-tau is frequently associated with other tau as well as self-associated. This image correlates with its secondary structure of 68% unstructured coil and 28.9% β-sheet. In contrast, the bovine tau in Figure 6C is not aggregated and has a monomer diameter of ~1.0 nm. This B tau preparation contains 5.1% α-helix, 23.1% β-turn, 7.4% β-sheet, and 66.4% unstructured coil. The B tau in Figure 6C and AD P-tau #1 in Figure 6A share a common tau monomer diameter of ~1.0 nm, whereas the HF AD P-tau chain has a diameter of only ~0.5 nm. None of these features is present in the control Pt-C metal film on the silver, which appears uniform with structures less than one-half to one-third the diameter of the Pt-C coated tau (Figs. 6D, 7C).

At high magnification (Fig. 7A) recombinant human tau, R-39, (19.7% β-sheet) was compared to B tau (~5.1% α-helix, ~23.1% β-turn, ~7.4% β-sheet). The R-39 tau chains have a diameter of ~0.5 nm. Some of the R-39 chains are self-aggregated and aggregated with other tau. These TEM observations qualitatively support R-39 tau's reported secondary structure estimate in Table I. B tau (Fig. 7B) contains monomer chains with a diameter of ~1.0 nm and a very low frequency of ~0.5 nm molecular chains and a few associated or tangled chains.

DISCUSSION AND CONCLUSIONS

Correlation of tau's secondary structure with TEM images

Human tau's largest isoform is 441 amino acids and can range in length from 32.5–112 nm, with its secondary structure displayed along its backbone (Ruben et al., 1991). In this section, we focus on the presence of α-helix, β-turns, and unstructured coil in tau samples prepared on silver filters and discuss β-sheet later. R-39 tau, HF AD P-tau, and AP AD P-tau when unassociated or entangled had chain diameters of ~0.5 nm (Figs. 6B, 7A, 9)—that of unordered coil. In contrast, AD P-tau #1 and bovine tau had diameters of ~1.0 nm (Figs. 6A,C, 7B). The α-helix content in these preparations was not sufficient (0 or 5%) to account for a ~1.0 nm chain diameter. With 36.8% and 23.1% β-turns and 57.1% and 64.3% coil, respectively, the ~1.0 nm diameter tau chains are due to a β-spiral structure (β-turns interspersed with coil) described previously (Ruben et al., 1991). Although AD P-tau #2 was isolated similarly to AD P-tau #1, this preparation was rinsed with alcohol and lyophilized. This treatment reduced its β-turn content to 10.4%—insufficient for β-spiral structure—and therefore had a chain diameter of ~0.5 nm.

Traditionally tau has been isolated by exposing each preparation to heat and acid pH treatment (Lindwall and Cole, 1984b). Tau isolated by this method has its phosphorylation state unchanged, is pure, and is able to stimulate microtubule assembly. Using heat- and acid-treated tau, we have been unable to reconstitute triple-stranded left-hand helical 2.1 nm tau polymers previously observed in neurofibrillary tangles and in bovine tau preparations (Ruben et al., 1991, 1992). Using a mixed microtubule associated protein (MAP) purification which does not require heat or acid treatment, we have observed 2.1 nm MAP polymers longitudinally associated with microtubules (Ruben et al., 1996). An isolation procedure that avoids heating and acid pH treatment should probably be able to produce 2.1 nm tau polymers.

AD P-tau is aggregated in solution

Apparently AD P-tau is aggregated in solution since it sediments between 27,000*g* to 200,000*g* and is pelleted, whereas normal tau in solution does not pellet at these forces (Alonso et al., 1994, 1996; Köpke et al., 1993). Reflecting its solution status, AD P-tau #1 and #2 were settled on silver filters in aggregated form and remained aggregated through the washing procedure (Figs. 5A, 6A, 8, 9). Its settled structures did not include PHF. Both AD P-tau #1 and #2 were more aggregated than any tau listed in Table I, and their secondary structures contained only 6.8 or 7.8% β-sheet. The CD deconvolution algorithm that we used was more accurate than one previously described (Schweers et al., 1994). The Lincomb method (Perczel et al., 1991) estimates the percentages of α-helix, β-turns, unstructured coil, β-sheet, and aromatics. The aromatics were then eliminated, and the nonzero protein secondary structures were increased accordingly in Table I. In a previously published estimate of tau's secondary structure by CD deconvolution, β-turns and aromatics were not taken into account (Schweers et al., 1994).

Abnormal phosphorylation correlates with aggregation, but the magnitude of β-sheet secondary structure does not

We found no correlation between AD P-tau aggregation and β-sheet content (Table I). There was, however,

a very strong correlation with abnormal levels of phosphorylation. Hydrofluoric acid appears to dephosphorylate AD P-tau completely (Greenberg et al., 1992). Samples were lyophilized after HF treatment at 4°C, resuspended in buffer for CD, and spread on silver filters. Unfiltered HF AD P-tau was also negatively stained but showed no detectable aggregation. The high β-sheet in HF AD P-tau #1 (33.8%) and #2 (26.0%) was likely due to the lyophilization of a preparation high in coil. The aggregation in both HF AD P-tau samples was ≤30% of the AD P-tau (Figs. 5B, 6B, 10). AD P-tau dephosphorylated with alkaline phosphatase (AP) to 90% completeness was reisolated from AP with heat and acid treatment and lyophilized, resuspended in 25 mM sodium phosphate, pH 6.5, for CD, and again spread on silver filters. The aggregation was reduced and was similar to HF AD P-tau: ≤40% of that of AD P-tau. AP AD P-tau treated with heat and acid, lyophilized, was 41.8% β-sheet and 58.2% unordered structure. Unphosphorylated recombinant human R-39 tau was isolated from *E. coli* with heating and acid treatment and was also lyophilized. Two samples resuspended in 25 mM sodium phosphate, pH 6.5, for CD and spread on silver filters were ≤20% aggregated. Together the samples averaged 19.7% β-sheet and 80.3% unordered structure. Bovine tau (B tau) with ~3.1 phosphates/tau, isolated without heat, acid treatment, or lyophilization, was composed of 7.4% β-sheet, similar to the AD P-tau #1 and #2, but was unaggregated.

All of the samples were isolated with dipositive and tripositive ion chelating agents (EDTA and EGTA) and were suspended in a 25 mM or 50 mM sodium phosphate, pH 6.5, buffer for CD and microscopy. Chelation should prevent or minimize the intermolecular connection of phosphates by di- or tripositive ion bridges. If this aggregation mechanism were important, B tau would have been aggregated with three phosphates, but it was not. Neither aggregation by β-sheet nor intermolecular bridging of phosphates by di- or tripositive ions can explain AD P-tau's aggregation.

The lack of correlation between β-sheet content and AD P-tau aggregation in PHF has also been reported. X-ray studies by Kirschner et al. (1986) suggested that dried PHF were associated by β-sheet, but their work was later contradicted convincingly by x-ray and fourier transform infrared spectroscopy (FTIR) work that indicated there was little β-sheet in hydrated PHF (Schweers et al., 1994). In addition, PHF have been disassembled by dephosphorylation with hydrofluoric acid (Greenberg et al., 1992) or with protein phosphatases, PP-2A and PP-2B (Wang et al., 1995). These studies also showed that tau's aggregation was related to abnormal phosphorylation and was absent after dephosphorylation.

AD P-tau is hydrophobic and aggregates hydrophobically

Serine (S) and threonine (T) residues are sites of phosphorylation in tau. These residues in the 441 amino acid human tau sequence (Goedert et al., 1989) are often close to positively charged lysine (pK 10.5), arginine (pK 12.5), or histidine (pK 6.0) residues. Positive charge neutralization by phosphates could shift tau's solution properties from hydrophilic to hydrophobic. There are 21 sites in AD P-tau where phosphates have been located (Brion et al., 1991; Hasegawa et al., 1992; Iqbal et al., 1989; Morishima-Kawashima et al., 1995a). A fraction of these phosphorylation sites is generally occupied on any single AD P-tau molecule (Köpke et al., 1993; Morishima-Kawashima et al., 1995b). Nonetheless, all of these sites were examined for their proximity to positively charged lysines (K), arginines (R), or histidines (H) in the full adult human tau sequence of 441 amino acids (Goedert et al., 1989). In Table IIA, phosphorylation sites closely associated with positive residue sites are listed, including Thr-217, which is within three residues of Arg-221 if an adjacent β-turn is taken into account (Ruben et al., 1991). Phosphate sites farther in sequence from basic amino acids in Table IIA are underlined. This does not preclude these phosphates being adjacent to positive residues in tau's β-spiral structure. Ser-262 is located in tau's first microtubule binding repeat and has been shown to greatly reduce tau's microtubule binding (Biernat et al., 1993).

Charge neutralization in AD P-tau would affect its normal isolation. Full-length tau contains more basic residues (44 lysines, 14 arginines, 10 histidines) than acidic residues (30 aspartate, 26 glutamate, 4 tyrosine) and has a basic protein isoelectric point (pI = 8.19). The excess basic residues make it possible to maintain tau in a perchloric acid solution (pH ~2.7) while other proteins precipitate—for example, MAP 2 (pI = 4.56). Occupation of five to nine phosphate sites appears sufficient to prevent AD P-tau from being solubilized by acidification. This has been confirmed by Alonso et al. (1994), who showed that acidification by $HClO_4$ (pH ~2.7) maintains normal tau in solution and does not solubilize Alzheimer's disease hyperphophorylated tau. Charge neutralization also explains why AD P-tau can be precipitated with 35% ammonium sulphate, whereas normal tau requires 45% ammonium sulphate (Köpke et al., 1993).

Positive charge neutralization of lower magnitude occurs in hyperphosphorylated fetal tau and normal adult tau

Fetal tau phosphates have been found on as many as 12 sites (Watanabe et al., 1993; Morishima-Kawashima et al., 1995b). Phosphorylated serines (S) and threonines (T) close to positive residues in fetal tau se-

TABLE II. Phosphorylation sites

A. Total number of human AD P-tau phosphorylation sites[1]

1	·····‖·····‖·····‖·····‖···K·S···‖	50
51	·····‖·····‖·····‖·····‖·····‖	100
101	·····‖·····‖··T··R···‖·····‖·····‖	150
151	·····‖·····‖·····‖·····‖·····S‖	200
201	·S·····SRSRT·S··T··‖·····RT··KS····‖·····‖	250
251	·····K‖·S·····‖·····‖·····‖····KS···‖	300
301	·····‖·····‖·····‖·····‖·····‖	350
351	·····‖·····‖·····‖·····‖····KS···S	400
401	··TS·RH·S‖·SS·····‖·S·····‖·····‖·	441

B. Total number of human fetal tau phosphorylation sites[1]

1	·····‖·····‖·····‖·····‖·····‖	50
51	·····‖·····‖·····‖·····‖·····‖	100
101	·····‖·····‖·····‖·····‖·····‖	150
151	·····‖·····‖·····KT·····‖·····SS‖	200
201	·S·····‖·····T··‖R·····RT··KS····‖·····‖	250
251	·····‖·····‖·····‖·····‖·····‖	300
301	·····‖·····‖·····‖·····‖·····‖	350
351	·····‖·····‖·····‖·····‖····KS···S	400
401	···S·RH·S‖··S·····‖·····‖·····‖·	441

C. Total number of normal rat tau phosphorylation sites[1]

1	·····‖·····‖·····‖·····‖·····‖	50
51	·····‖·····‖·····‖·····‖·····‖	100
101	·····‖·····‖·····‖·····‖·····‖	150
151	·····‖·····‖KT·····‖·····S··S·····‖	200
201	·····‖·····‖RT·····‖·····‖·····‖	250
251	·····‖·····‖·····‖·····‖·····‖	300
301	·····‖·····‖·····‖·····‖·····‖	350
351	·····‖·····‖·····‖·····‖····S·R··‖	400
401	·····‖·····‖·····‖·····‖·····‖	441

[1]Serine (S) and threonine (T) residues found to be phosphorylated and located close to basic residues in the 441 amino acid sequence (lysine, K; arginine, R; histidine, H) as well as *S* and *T* not close in sequence to positive charges.

quence are in Table IIB. Phosphates not close in sequence to positive residues included in Table IIB are underlined. Fetal tau is able to stimulate microtubule assembly (Yoshida and Ihara, 1993), whereas AD P-tau is unable to stimulate microtubule formation (Alonso et al., 1994; Iqbal et al., 1986, 1994). In adult rat tau a total of five sites were found to be phosphorylated and adjacent to positive charges in Table IIC, and two (underlined) were not. Mapping of phosphates near positive residues suggests that AD P-tau has more of these pairings with potentially greater charge neutralization.

The inverse temperature transition of all tau preparations occurs at a higher temperature than that of AD P-tau, indicating that it is more hydrophobic

The coil circular dichroism theta value at 197 nm clearly increases with increasing temperature to a characteristic plateau value. It has been suggested that the clathrate water structure around hydrophobic groups collapses as tau shortens, bringing the groups closer together as in the elastic protein elastin (Urrey, 1990, 1993). This transition is described as the inverse temperature transition, since random coil is normally produced when a protein is heated from 20°C to 80°C. The inverse temperature transition, T_t, is defined as the temperature at half-maximum increase at 197 nm in the CD spectrum. The T_t for dephosphorylated HF AD P-tau was ~37.5°C and for unphosphorylated R-39 tau was ~38°C. For conventionally isolated bovine tau with ~3.1 phosphates per monomer, T_t was ~32°C (Ruben et al., 1991). The T_t for AD P-tau, with five to nine phosphates/monomer (AD P-tau #1, 50 mM sodium phosphate, pH 6.5) was measured at ~20°C (24.5°C, T_t equivalent in 25 mM sodium phosphate) and ~28°C (AD P-tau #2, 25 mM sodium phosphate, pH 6.5). According to Urry (1993), the T_t of an uncharged elastic protein is expected to increase with phosphorylation and decrease with dephosphorylation. The inverse temperature transition in AD P-tau decreased from 38°C to 24.5°C or 28°C, confirming the occurrence of charge neutralization.

Conclusions

Tau isolated by heat and acid treatment retains its phosphates and can stimulate microtubule assembly. The secondary structure of this tau is predominately unstructured coil with a chain diameter of ~0.5 nm. We isolated two tau samples, avoiding acid and heat treatments, and produced a protein with a diameter of ~1.0 nm, 23.1% or 36.8% β-turns, and the remainder mostly coil. The 1.0 nm diameter is due to a β-spiral structure (Ruben et al., 1991) since tau's α-helix content is too low to produce more than 5% of the monomer with a 1.0 nm diameter. The β-spiral structure is likely present in triple-stranded 2.1 nm tau polymer that has been visualized longitudinally associated with microtubules

(Ruben et al., 1996). When tau is mainly coil we have been unable to produce 2.1 nm tau polymer, suggesting that a monomer secondary structure with an ~1.0 nm diameter is important for 2.1 nm tau polymer formation. β-spiral secondary structure occurs in elastic proteins which shorten with increasing temperature and that undergo an inverse temperature transition (Urry, 1993). Our results show that an inverse temperature transition occurs in tau with either a native or a denatured β-spiral secondary structure (AD P-tau #1 and #2 in Fig. 2e). The inverse temperature transition is the mechanism by which native or denatured tau (~3.1 phosphates) goes from an extended to a shortened state at 37°C and stimulates microtubule assembly (Ruben et al., 1991).

Until now no mechanism for the formation of neurofibrillary tangles of AD P-tau in Alzheimer's disease tau had been identified. We have presented evidence that neither total β-sheet content nor a positive ion intermolecular bridging of phosphates is responsible for AD P-tau aggregation. Instead we suggest a hydrophobic mechanism for aggregation based on three lines of evidence. First, many phosphorylation sites on AD P-tau occur adjacent to positive residues (~13 of 21) just as they do in fetal tau (~7 of 12) and in adult rat tau (~3 of 5), but more charges are neutralized in AD P-tau. Second, although tau is normally soluble in 2.5% perchloric acid, AD P-tau is not (Alonso et al., 1994). AD P-tau can be precipitated from 35% ammonium sulfate, whereas normal tau is precipitated only from 45% ammonium sulfate (Köpke et al., 1993). Finally, the inverse temperature transition, T_t, of AD P-tau is strongly reduced below body temperature (37°C) supporting charge neutralization. A lowered inverse temperature transition (24.5 and 28°C) indicates that AD P-tau has properties of an elastic hydrophobic protein that coalesces and precipitates from solution below body temperature at 24–29°C (Urry, 1993).

ACKNOWLEDGMENTS

These studies were supported by NIH grants AG11054 (G.C. Ruben, Assistant), (I. Grundke-Iqbal), and AG05892, AG08070, and TW00507 (K. Iqbal). The authors thank Thomas B. Roos, Ph.D., for critically reading the manuscript. G.C.R. thanks the Rippel EM lab for the use of its electron microscopes.

REFERENCES

Alonso, A. del C., Zaidi, T., Grundke-Iqbal, I., and Iqbal, K. (1994) Role of abnormally phosphorylated tau in the breakdown of microtubules in Alzheimer disease. Proc. Natl. Acad. Sci. U.S.A., 91:5562–5566.

Alonso, A. del C., Grundke-Iqbal, I., and Iqbal, K. (1996) Alzheimer's disease hyperphosphorylated tau sequesters normal tau into tangles of filaments and disassembles microtubules. Nature Medicine, 2:1–5.

Bancher, C., Brunner, C., Lassmann, H., Budka, H., Jellinger, K., Wiche, G., Seitelberger, F., Grundke-Iqbal, I., Iqbal, K., and Wisniewski, H.M. (1989) Accumulation of abnormally phosphorylated tau precedes the formation of neurofibrillary tangles in Alzheimer's disease. Brain Res., 477:90–99.

Barcikowska, M., Wisniewski, H.M., Bancher, C., and Grundke-Iqbal, I. (1989) About the presence of paired helical filaments in dystrophic neurites participating in the plaque formation. Acta Neuropathol. (Berl.), 78:225–231.

Bensadoun, A., and Weinstein, D. (1976) Assay of proteins in the presence of interfering materials. Anal. Biochem., 70:241–250.

Biernat, J., Gustke, N., Drewes, G., Mandelkow, E.-M., and Mandelkow, E. (1993) Phosphorylation of Ser-262 strongly reduces binding of tau to microtubules: Distinction between PHF-like immunoreactivity and microtubule binding. Neuron, 11:153–163.

Brion, J.P., Hager, D.P., Bruce, M.T., Couck, A.M., Flament-Durant, J., and Anderton, B.T. (1991) Tau in Alzheimers neurofibrillary tangles. Biochem. J., 273:127–133.

Ciardelli, T.L., Landgraf, B., Gadski, R., Strand, J., Cohen, F.E., and Smith, K.A. (1988) A design approach to the structural analysis of interleukin-2. J. Mol. Recognit., 1:42–47.

Dickson, D.W., Farlo, J., Davies, P., Crystal, H., Fuld, P., and Yen, S.C. (1988) Alzheimer's disease. Am. J. Pathol., 132:86–101.

Francis, P.T., Chessell, I.P., Webster, M.-T., Clarke, N.A., Procter, A.W., Alder, J.T., Chen, C., Qume, M., Zeman, S., Dijk, S., and Bowen, D. (1995) Is improvement of cholinergic transmission the correct strategy for Alzheimer's disease? In: Research Advances in Alzheimer's Disease and Related Disorders. K. Iqbal, J.A. Mortimer, B. Winblad, and H.M. Wisniewski, eds. J. Wiley & Sons, New York, pp. 273–281.

Goedert, M., Spillantini, M.G., Jakes, R., Rutherford, D., and Crowther, R.A. (1989) Multiple isoforms of human microtubule-associated protein tau: Sequences and localization in neurofibrillary tangles of Alzheimer's disease. Neuron, 3:519–526.

Goedert, M., Spillantini, M.G., Cairns, N.J., and Crowther, R.A. (1992) Tau proteins of Alzheimer paired helical filaments: Abnormal phosphorylation of all six brain isoforms. Neuron, 8:159–168.

Greenberg, S.G., Davies, P., Schein, J.D., and Binder, L.I. (1992) Hydrofluoric acid-treated τ_{PHF} proteins display the same biochemical properties as normal τ^*. J. Biol. Chem., 267:564–569.

Grundke-Iqbal, I., Iqbal, K., Quinlan, M., Tung, Y.-C., Zaidi, M.S., and Wisniewski, H.M. (1986a) Microtubule-associated protein tau: A component of Alzheimer paired helical filaments. J. Biol. Chem., 261:6084–6089.

Grundke-Iqbal, I., Iqbal, K., Tung, Y.-C., Quinlan, M., Wisniewski, H.M., and Binder, L. I. (1986b) Abnormal phosphorylation of the microtubule associated protein tau in Alzheimer cytoskeletal pathology. Proc. Natl. Acad. Sci. U. S. A. 83:4913–4917.

Hasegawa, M., Morishima-Kawashima, M., Takio, K., Suzuki, M., Titani, K., and Ihara, Y. (1992) Protein sequence and mass spectrmetric analyses of tau in Alzheimer disease brain. J. Biol. Chem., 267:17047–17054.

Iqbal, K., and Grundke-Iqbal, I. (1995) Alzheimer abnormally phosphorylated tau is more hyperphosphorylated than the fetal tau and causes the disruption of microtubules. Neurobiol. Aging, 16:375–379.

Iqbal, K., Grundke-Iqbal, I., Zaidi, T., Merz, P.A., Wen, G.Y., Shaikh, S.S., Wisniewski, H.M., Alafuzoff, I., and Winblad, B. (1986) Defective brain microtubule assembly in Alzheimer's disease. Lancet, 2:421–426.

Iqbal, K., Grundke-Iqbal, I., Smith, A.J., George, L., Tung, Y.-C., and Zaidi, T. (1989) Identification and localization of a tau peptide to paired helical filaments of Alzheimer disease. Proc. Natl. Acad. Sci. U. S. A., 86:5646–5650.

Iqbal, K., Zaidi, T., Bancher, C., and Grundke-Iqbal, I. (1994) Alzheimer paired helical filaments: Restoration of the biological activity by dephosphorylation. FEBS Lett., 349:104–108.

Jakes, R., Novak, M., Davison, M., and Wischik, C.M. (1991) Identification of 3- and 4-repeat tau isoforms within the PHF in Alzheimer's disease. EMBO J., 10:2725–2729.

Katzman, R., Terry, R.D., Deteresa, R., Brown, R., Davies, P., Fuld, P., Renling, X., and Peck, A. (1988) Clinical, pathologic and neurochemical changes in dementia: A subgroup with preserved mental status and numerous neocortical plaques. Ann. Neurol., 23:138–144.

Kidd, M. (1964) Alzheimer's disease. An electron microscopical study. Brain, 87:307–320.

Kirschner, D.A., Abraham, C., and Selkoe, D.J. (1986) X-ray diffraction from intraneuronal paired helical filaments and extraneuronal amyloid fibers in Alzheimer disease indicates cross-β conformation. Proc. Natl. Acad. Sci. U. S. A., 83:503–507.

Köpke, E., Tung, Y.-C., Shaikh, S., Alanso, A. del C., Iqbal, K., and Grundke-Iqbal, I. (1993) Microtubule-associated protein tau: Abnormal phosphorylation of a non-paired helical filament pool in Alzheimer's disease. J. Biol. Chem., 268:24374–24384.

Lindwall, G., and Cole, R.D. (1984a) Phosphorylation affects the ability of the tau protein to promote microtubule assembly. J. Biol. Chem., 259:5301–5305.

Lindwall, G., and Cole, R.D. (1984b) Purification of tau protein and the

occurrence of two phosphorylation states of tau in brain. J. Biol. Chem., 258:12211–12215.
Morishima-Kawashima, M., Hasegawa, M., Takio, K., Suzuki, M., Yoshida, H., Watanabe, A., Titani, K., and Ihara, Y. (1995a) Hyperphosphorylation of tau in PHF. Neurobiol. Aging, 16:365–380.
Morishima-Kawashima, M., Hasegawa, M., Takio, K., Suzuki, M., Yoshida, H., Watanabe, A., Titani, K., and Ihara, Y. (1995b) Proline-directed and non-proline-directed phosphorylation of PHF-tau. J. Biol. Chem., 270:823–829.
Perczel, A., Hollosi, M., Tusnady, G., and Fasman, G.D. (1991) Convex constraint analysis: A natural deconvolution of circular dichroism spectra. Protein Eng., 4:669–679.
Ruben, G.C., Harris, Jr., E.D., and Nagase, H. (1988) Electron microscope studies of free and proteinase-bound duck ovostatins (ovomacroglobulin I): mode of ovostatin structure and its transformation upon proteolysis. J. Biol. Chem., 263:2861–2869.
Ruben, G.C. (1989) Platinum (Pt-C) vertically shadowed Pt-C replicas for imaging individual molecules in freeze-etched biological DNA and material science metal and plastic specimens. J. Electr. Microsc. Tech., 13:335–351.
Ruben, G.C., Iqbal, K., Grundke-Iqbal, I., Wisniewski, H.M., Ciardelli, T.L., and Johnson, J.E., Jr. (1991) The microtubule associated protein tau forms a triple-stranded left-hand helical polymer. J. Biol. Chem., 266:22019–22027.
Ruben, G.C., Iqbal, K., Wisniewski, H.M., Johnson, J.E., Jr., and Grundke-Iqbal, I. (1992) Alzheimer neurofibrillary tangles contain 2.1 nm filaments structurally identical to the microtubule associated protein tau. Ultrahigh resolution transmission electron microscopy study of large and senile plaque core amyloid. Brain Res., 590:164–170.
Ruben, G.C., Iqbal, K., Grundke-Iqbal, I., and Johnson, J.E., Jr. (1993) The organization of the microtubule associated protein tau in Alzheimer paired helical filaments. Brain Res., 602:1–13.
Ruben, G.C., Novak, M., Fasulo, C.C., and Iqbal, K. (1995) Alzheimer paired helical filaments, untreated and pronase digested, studied by vertical platinum-coated replication and high resolution transmission electron microscopy. Brain Res., 675:1–12.
Ruben, G.C., Abuza, A., and C., Grundke-Iqbal, I., and Iqbal, K. (1996) Taxol stabilized rat brain microtubules with microtubule-associated proteins (MAP) studied by vertically platinum-carbon (Pt-C) replicated: New [illegible] resolution images for evaluating the relationship of MAPs to microtubules. Neuroscience-Net, 1:article 10002.
Schweers, O., Schonbrunn-Hanebeck, E., Marx, A., and Mandelkow, E. (1994) Structural studies of tau protein and Alzheimer paired helical filaments show abundance of beta-structure. J. Biol. Chem., 269:24290–24297.
Singh, T.J., Haque, N., Grundke-Iqbal, I., and Iqbal, K. (1995) Rapid Alzheimer-like phosphorylation of tau by the synergistic actions of non-proline-dependent protein kinase and GSK-3. FEBS Lett., 358:267–272.
Urry, D.W. (1990) Protein folding and assembly: An hydration-mediated free energy transduction. In: Protein Folding: Deciphering the Second Half of the Genetic Code, L.M. Gierasch and J. King, eds. American Association for the Advancement of Science, Washington, DC, pp. 63–77.
Urry, D.W. (1993) Molecular machines: How motion and other functions of living organisms can result from reversible chemical changes. Angew. Chem. Int. Ed. Engl., 32:819–841.
Wang, J.-Z., Gong, C.-X., Zaidi, T., Grundke-Iqbal, I., and Iqbal, K. (1995) Dephosphorylation of Alzheimer paired helical filaments by protein phosphatase-2A and -2B. J. Biol. Chem., 270:4854–4860.
Watanabe, A., Hasegawa, M., Suzuki, M., Takio, K., Morishima-Kawashima, M., Titani, K., Arai, T., Kosik, K.A., and Ihara, Y. (1993) In vivo phosphorylation sites in fetal and adult rat tau. J. Biol. Chem., 268:25712–25717.
Yoshida, H., and Ihara, Y. (1993) Tau in paired helical filaments is functionally distinct from fetal tau: Assembly incompetence of paired helical filament-tau. J. Neurochem., 61:1183–1186.

APPENDIX 1
Stereoscopic TEM images of AD P-tau, HF AD P-tau, R3.0 tau, and 3 tau on silver filters

Stereoscopic images of the tau preparations on the silver filter are an essential part of the evidence evaluating AD P-tau aggregation with and without dephosphorylation with HF or alkaline phosphatase.

Figure 8 contains two stereo images of AD P-tau #1 showing networks of ridges. In Figure 8A a network of condensed AD P-tau ridges appears on top of a silver hill and in the depression behind it. The network ridges appear compact, with the long axis of AD P-tau monomers protruding from the ridges. Adjacent to the ridges are other mostly horizontal tau molecules adhering to the silver surface. At higher magnification (Fig. 8B), the stereo images show similar networked ridges of condensed AD P-tau. The AD P-tau monomers in the ridges protrude from them, as in Figure 6A. In the depression behind the silver hill, sections of the network of ridges show open gaps. It is doubtful that the gaps are an intrinsic property of the network; rather, the AD P-tau in the ridges may be separated in the spreading process or removed in the process of washing. Another image feature is the presence of short or punctate molecules on the surface. In stereoscopic imaging these molecules frequently appear with their long axis pointing toward the observer. Without stereo observation it is difficult to know whether the tau molecules are short pieces of degraded tau or extended molecules standing on their long axis.

Stereo images of AD P-tau #2 are shown in Figure 9. AD-P tau #2 molecules are generally thinner than the AD P-tau #1 in Figure 8 and are often not distinguishable as individuals in their association with other AD P-tau. In Figure 8A there are film-like structures of AD P-tau at the center and bottom of the image. In the upper left of this image is a network of associated AD P-tau, and in the upper center is a thin filament. To the right are small groupings of condensed AD P-tau. These AD P-tau structures were not evident in Figure 8, where the AD P-tau was deposited from a solution containing approximately five times higher concentration. We suspect that these structures may have been present in AD P-tau sample #1 but masked by the network/film conformation predominant at higher concentrations (Fig. 5A). In Figure 9B, a 2–8 nm wide filament extends above the surface, branching from an irregular filament or surface ~16 nm × ~108 nm. There are other filaments present, but none of them have the characteristic size or helicity of PHF. Finally, Figure 9C displays most of the possible forms of aggregation. At the upper right is a large sheet standing on its side, at the left are small groups of condensed AD P-tau forming particles, and at the bottom of this image are two small filaments. Neither of these filaments resembles PHF nor shows helical conformation. The aggregated AD P-tau structures in Figures 8 and 9 represent the full range of the structures observed in the AD P-tau preparations.

Nonenzymatic hydrofluoric acid treatment of AD P-tau produces ~100% dephosphorylation (Greenberg et al., 1992), whereas the alkaline phosphatase treatment used in Table I has only ~90% effectiveness (unpublished observation). Figure 10 shows two stereo

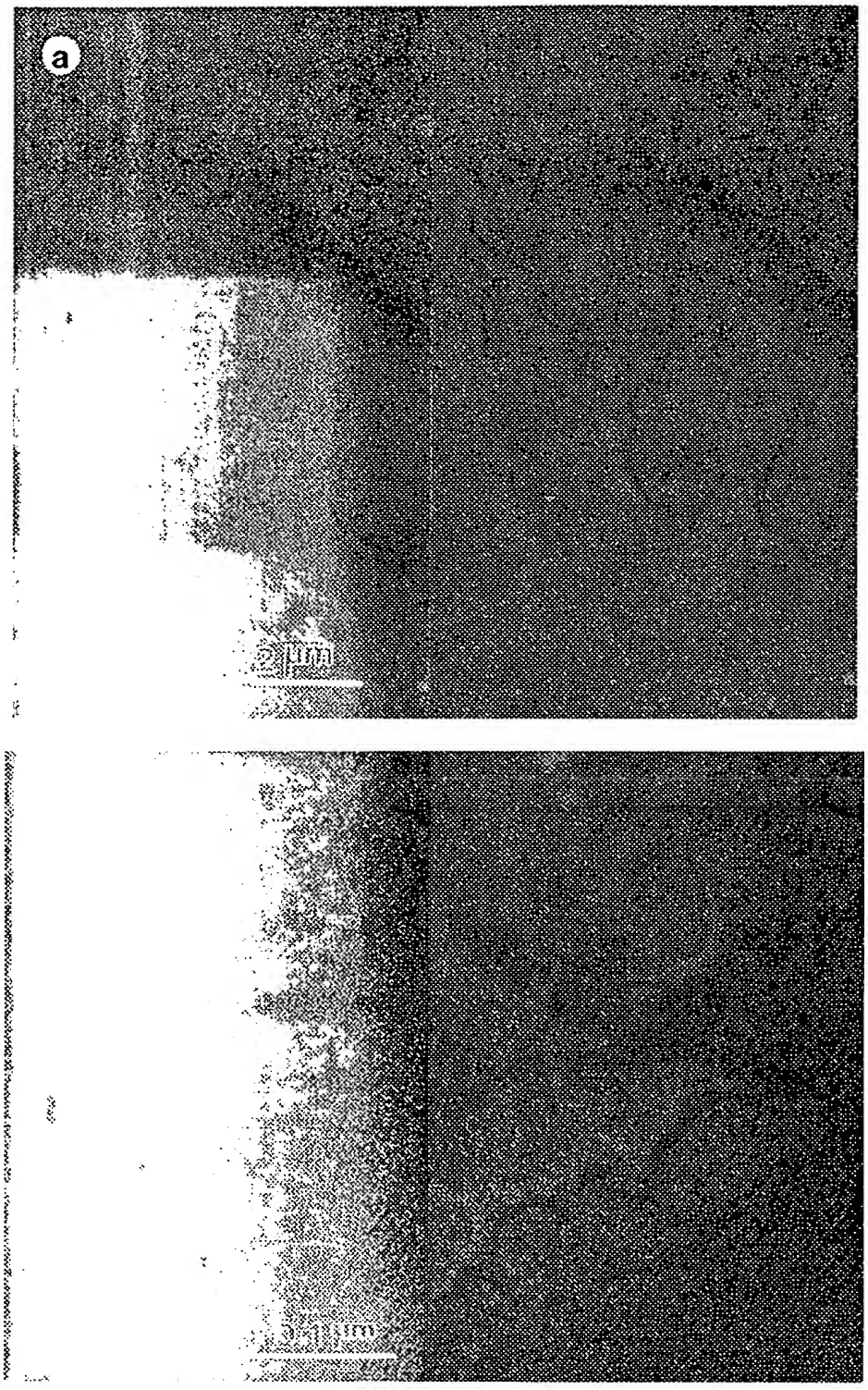


Fig. 8. Stereo i[mage] ...ried, vertically replicated AD P-tau #1 (~160 ... ...et heat, acid, or lyophilization treatment in 50 m... ...ed on a 0.2 μm silver filter, washed, and fro... ... This sample was vertically replicated with 0... ... carbon coated with 12.1 nm. This is the same ... ...n Figs. 5A and 6A. A: Stereo image, 15° tilt an... ...ns interconnected condensed AD P-tau forms a ... ...work which contains variable width ridges rang... ...'s network is associated with the filter's silver s... ...not disassemble or remove it. Individual tau molecules are also associated with the filter's silver surface between the connectors. ×150,000. **B:** This stereo image (5° tilt angle between images) shows the same high magnification section of condensed AD P-tau as in Fig. 6A. The AD P-tau monomer chains (~1.0 nm in diameter) appear to be associated longitudinally and stand on the filter's silver surface with their monomer ends protruding from the ridge's surface. Individual AD P-tau monomers also appear to lie on the surface between the condensed AD P-tau network's connecting ridges. ×230,000.

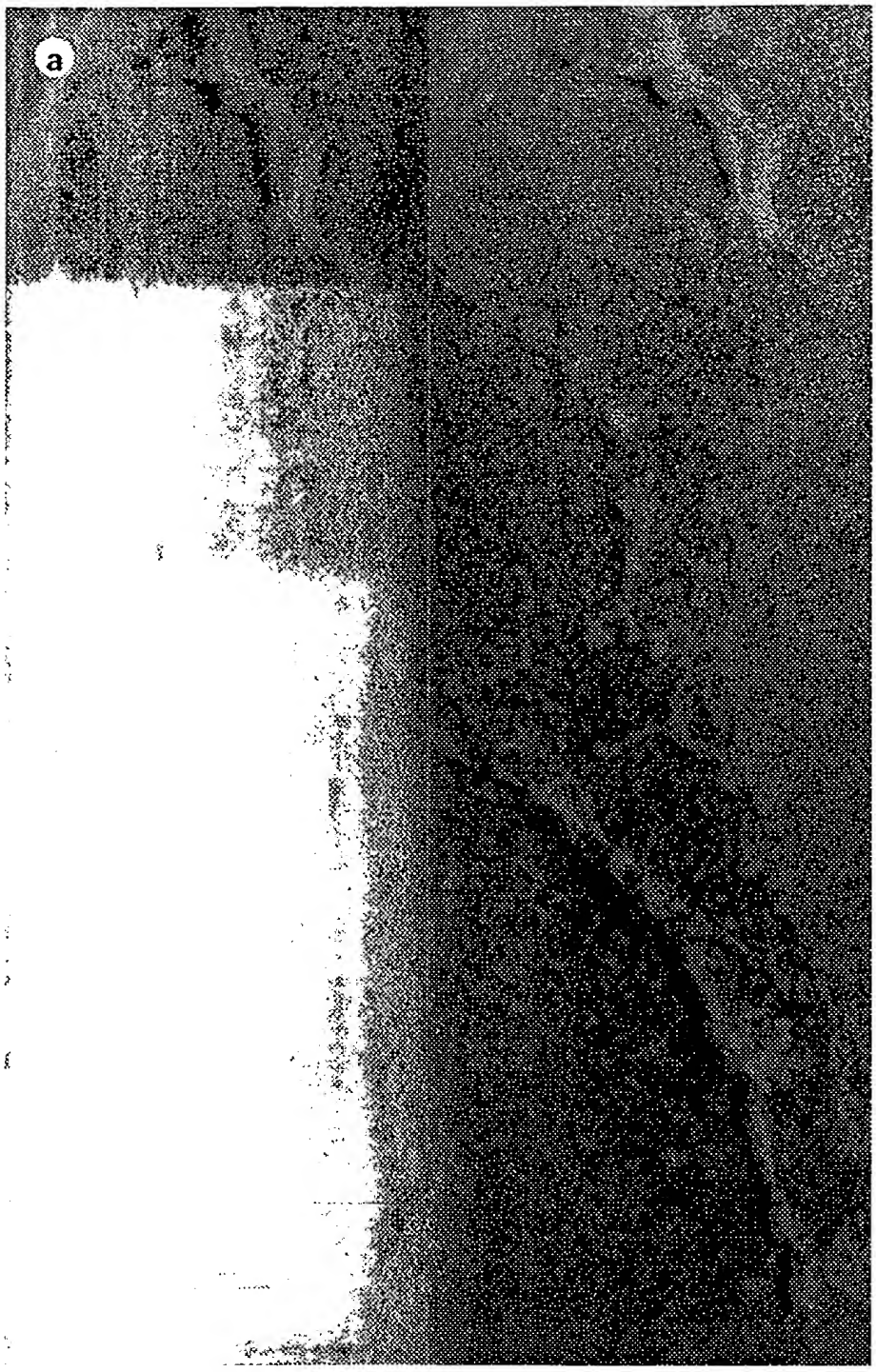


Fig. 9. Stereo [illegible] vertically replicated AD P-tau #2 (~35 µg [illegible] treatment. This sample was precip[illegible] lyophilized. This lyophilized AD P-[illegible] 5 mM NaH_2PO_4 (pH 6.5) and spread [illegible], and frozen for freeze-drying. Thi[illegible] with 0.98 nm Pt-C and rotary carbon [illegible] image, 15° tilt angle between images. T[illegible] structures composed of AD P-tau appear [illegible] in the middle and lower section of this im[illegible] section of this image is a network of variabi[illegible] as nearly as the ridges of AD P-tau #1 in Fi[illegible] regularity in these network structures is evid[illegible] filament can be seen that is 4–5 nm wide and ~[illegible] of center in the upper third of the image are s[illegible] AD P-tau that range from 8–20 nm in diam[illegible] in this AD P-tau preparation. ×232,000. **B:** This stereo image (10° tilt angle between images) shows a ribbed filament, 2–8 nm wide, extending towards the viewer. At its lower end it branches from a filament ~16 nm wide and ~108 nm long of no discernible regularity. Below this structure is a sheet-like structure on its side as well as irregular filaments 3–8 nm in width. At the left is another irregular branching filament. ×250,000. **C:** This stereo image (10° tilt angle between images) contains a large, sheet-like AD P-tau structure fringed with a transparent carbon film in the upper right. In the upper left quadrant, small clumps of condensed AD P-tau form particles, 8–20 nm in diameter, that are also surrounded by a carbon film halo which was excluded from the measurement. In the lower right quadrant of this stereo pair are two filaments: one is 6.6–10 nm wide by ~99 nm long, and the second at the right is 7–13 nm wide and ~120 nm long. Neither of these filaments appears helical like the PHF. ×151,000.

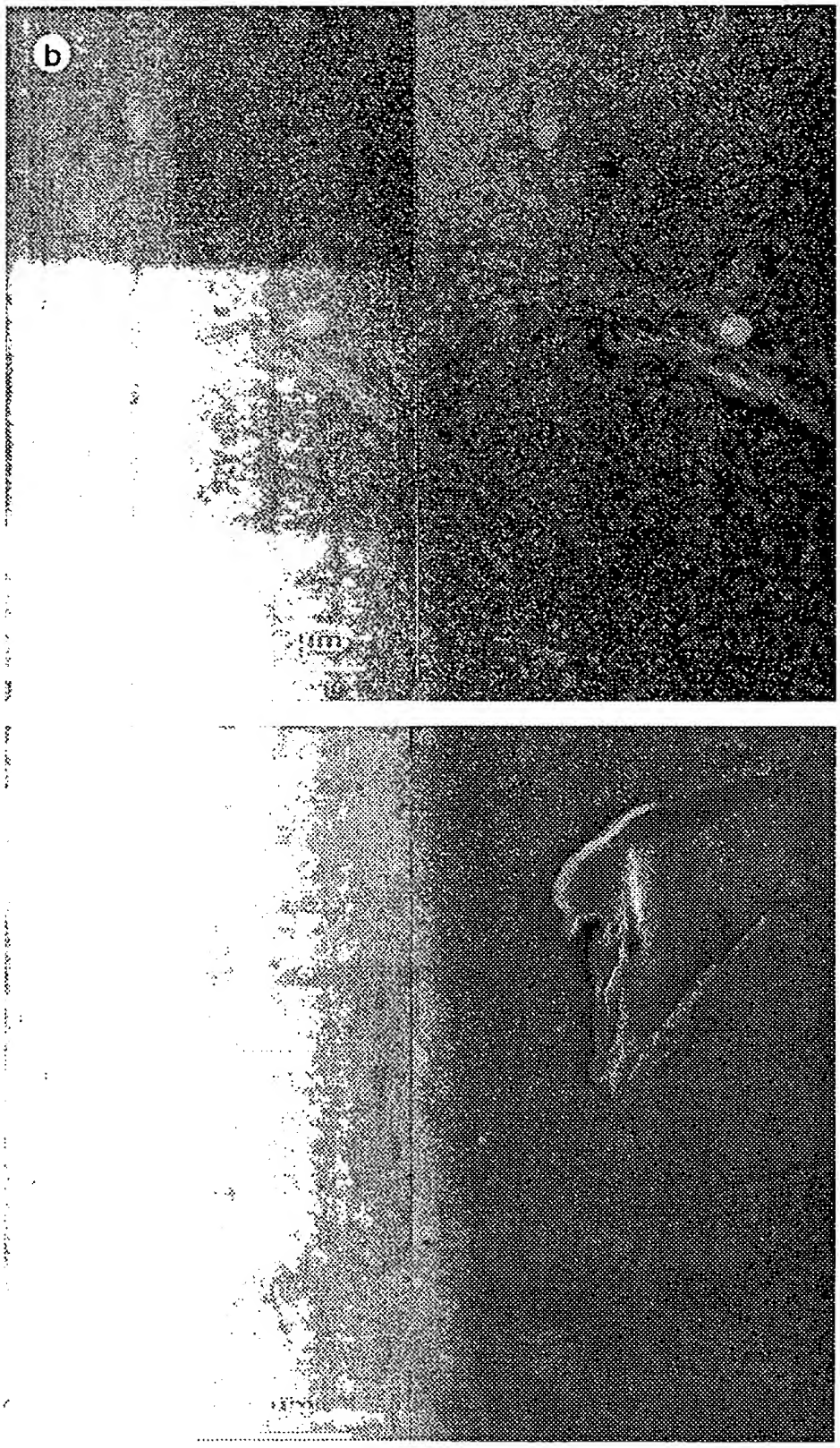


Figure 9 (Continued.)

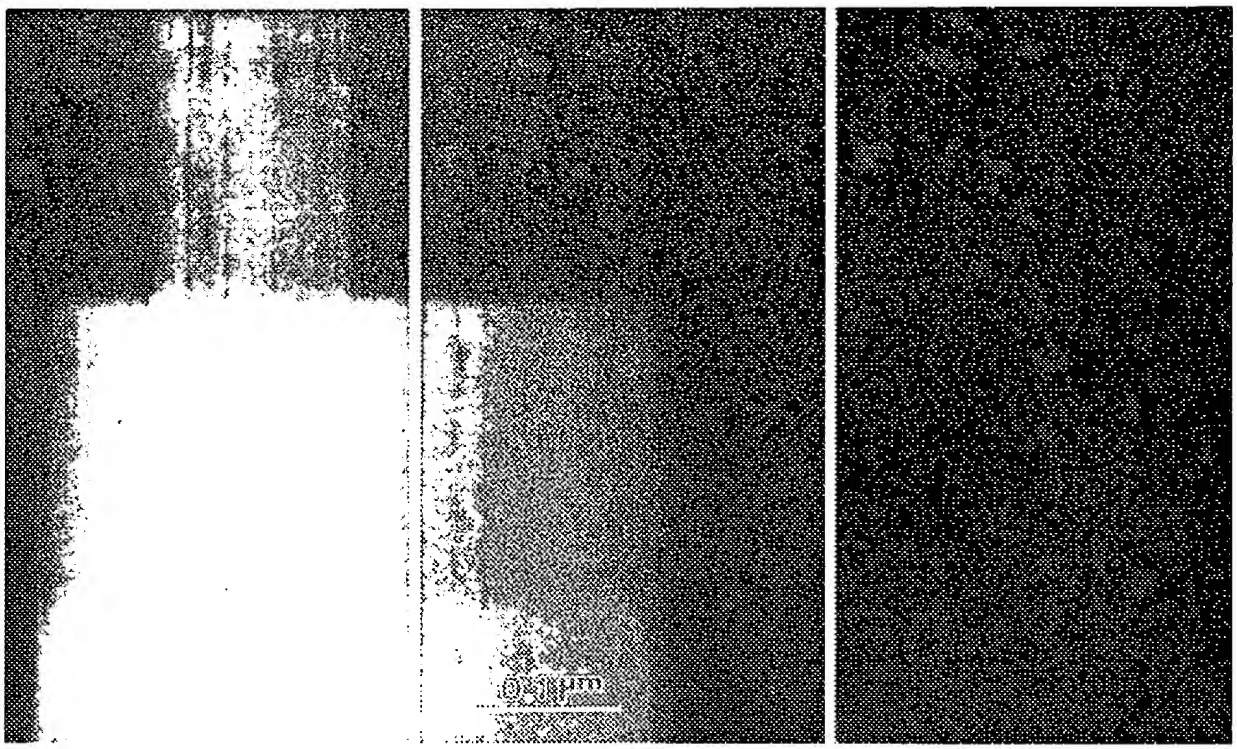


Fig. 10. ...ca'v replicated HF AD P-tau f... 116.5) and spread on a 0.2 ... ze-drying. This samp le w... ...nd rotary carbon coate... ...ion with HF at 4°C, we did ... ...ures s... ...n in Figs. 6A, 8, and 9. T... ...ed of three TEM tilt images of +10°, +5°, and −10°. In the left-hand pair of images, the tau chains on the surface appear to be mostly separate with a few associated (see lower half of image in both stereo pairs). Some of the small clumps of tau rising above the surface do not appear associated, but the majority are associated. An occasional filament composed of tau monomers can also be found in the HF AD P-tau, but their frequency is low compared to the AD P-tau #2.

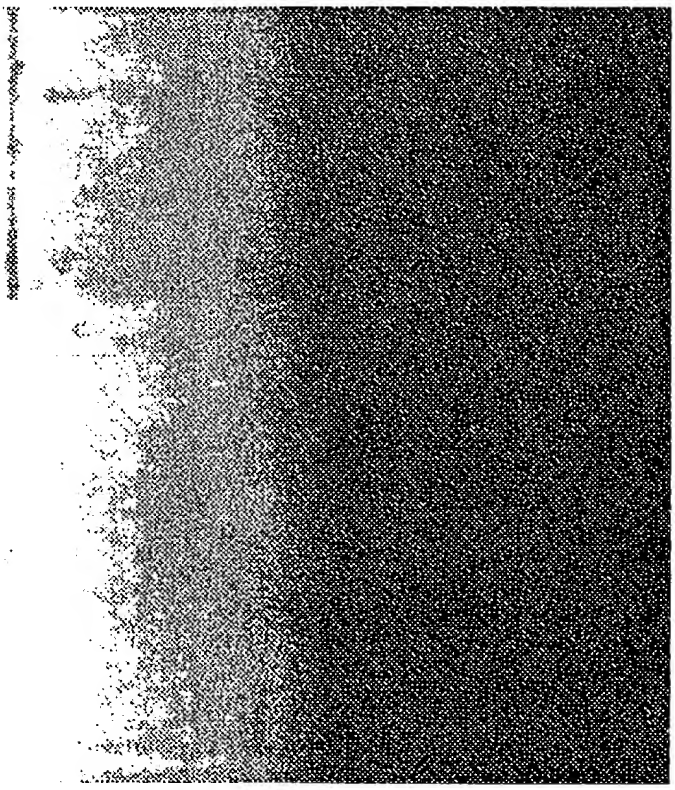


Fig. 11. ...seen images) of freeze-dried, vertic... ...), suspended in 25 mM NaHP... ...ter, washed, and frozen for ... ...pliated with 1.0 nm Pt-C an... ...R-39 tau covered surface appe... ...on the surface as well as a few clumped R-39 tau standing on the surface surrounded by a carbon halo. These tau associations are similar to those seen in the HF AD P-tau. They are also less frequent, and no filamentous structures of tau monomer have been detected in this preparation. ×235,000.

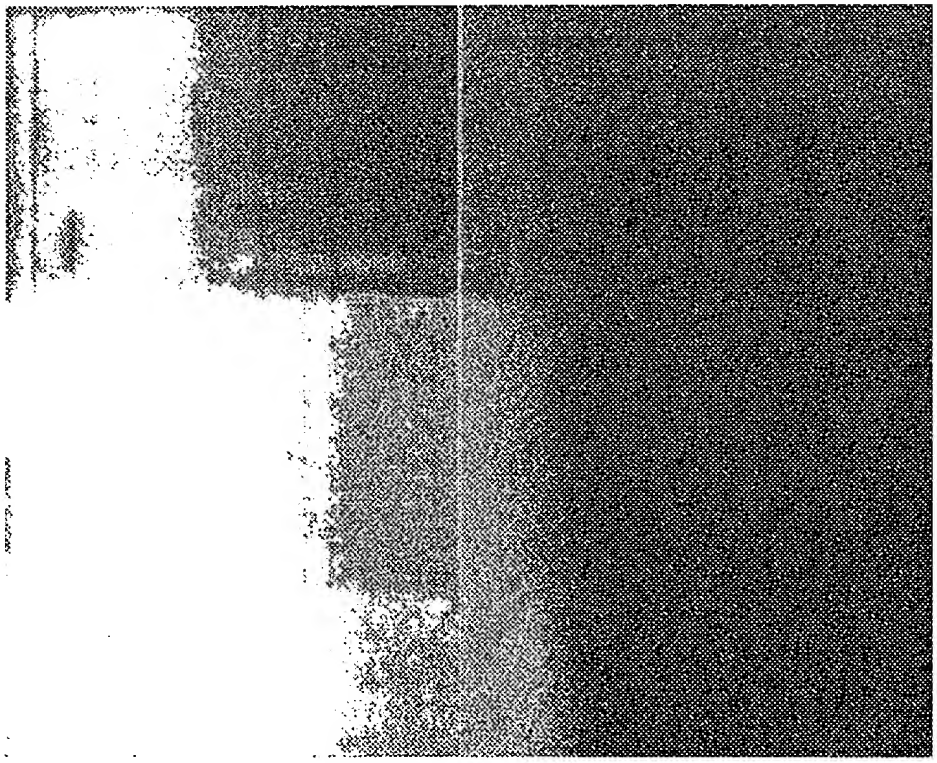


Fig. 12. Stereo [illegible] images) of freeze-dried, vertically r[illegible] 2 mM $NaHPO_4$ (pH 6.5) spread [illegible], and frozen for freeze-drying. Thi[illegible] 0.98 nm Pt-C and rotary carbo[illegible] -39 tau in Fig. 7A, bovine tau w[illegible] or lyophilization. Its secondary structure contains β-turns and α-helix, which explains why the tau diameters in this image appear larger those of HF AD P-tau or R-39 tau, which are mostly coil (see Figs. 6, 7). Although the bovine tau monomers lie close to each other, there appears to be no apparent association between them. ×221,000.

pairs in which [illegible]-tau are present but networks [illegible] 5.', 8, 9A) are absent. Filam[illegible] only rarely present and w[illegible] a secondary structure of 28.9% [illegible] is sparsely aggregated, unlike [illegible] aggregated AD-P tau samples whic[illegible] β-sheet.

Figure 11 [illegible] human tau, R-39. Except for a f[illegible] appear to form particles, this [illegible] does not appear to be aggregated, although we know it is in high magnification images. This sample also has more β-sheet than AD P-tau but was considerably less aggregated than AD P-tau.

Finally, in Figure 12 we have stereo images of B tau isolated without heat or acid treatment. The B tau monomers generally appear to have larger chain diameters than R-39 tau or HF AD P-tau. This preparation has roughly the same amount of β-sheet (7.4%) as AD P-tau and is unaggregated.

JOURNAL OF VIROLOGY, July 1996, p. 4623–4631
0022-538X/96/$04.00+0

Vol. 70, No. 7

Assemblons: Nuclear Structures Defined by Aggregation of Immature Capsids and Some Tegument Proteins of Herpes Simplex Virus 1

PATRICIA L. WARD, WILLIAM O. OGLE, AND BERNARD ROIZMAN*

The Marjorie B. Kovler Viral Oncology Laboratories, The University of Chicago, Chicago, Illinois 60637

Received 5 February 1996/Accepted 25 March 1996

In cells infected with herpes simplex virus 1 (HSV-1), the viral proteins ICP5 (infected-cell protein 5) and VP19c (the product of U_L38) are associated with mature capsids, whereas the same proteins, along with ICP35, are components of immature capsids. Here we report that ICP35, ICP5, and U_L38 (VP19c) coalesce at late times postinfection and form antigenically dense structures located at the periphery of nuclei, close to but not abutting nuclear membranes. These structures were formed in cells infected with a virus carrying a temperature-sensitive mutation in the U_L15 gene at nonpermissive temperatures. Since at these temperatures viral DNA is made but not packaged, these structures must contain the proteins for immature-capsid assembly and were therefore designated assemblons. These assemblons are located at the periphery of a diffuse structure composed of proteins involved in DNA synthesis. This structure overlaps only minimally with the assemblons. In contrast, tegument proteins were located in asymmetrically distributed structures also partially overlapping with assemblons but frequently located nearer to nuclear membranes. Of particular interest is the finding that the U_L15 protein colocalized with the proteins associated with viral DNA synthesis rather than with assemblons, suggesting that the association with DNA may take place during its synthesis and precedes the involvement of this protein in packaging of the viral DNA into capsids. The formation of three different compartments consisting of proteins involved in viral DNA synthesis, the capsid proteins, and tegument proteins suggests that there exists a viral machinery which enables aggregation and coalescence of specific viral protein groups on the basis of their function.

The herpes simplex virus 1 and 2 (HSV-1 and HSV-2) particles consist of four concentric structural elements, i.e., a central DNA core; a capsid consisting of the products of the genes U_L18 (VP23), U_L19 (VP5), U_L26 (VP21, VP22a, and VP24), U_L35 (VP26), and U_L38 (VP19c); an amorphous protein structure called the tegument which surrounds the capsid; and an envelope containing several glycoproteins (4, 6, 8, 14, 21, 22, 30, 34, 36). The HSV genomes are both transcribed and replicated in the nucleus, which is also the site of immature viral capsid assembly. The studies summarized in this report were initiated following the observation that the product of a newly discovered open reading frame designated $U_L43.5$ and mapping antisense to U_L43 colocalized with capsid proteins in dense structures located at the periphery of the nucleus late in infection (44). Spurred by these observations, we began a systematic study of the localization of several classes of structural and nonstructural viral proteins during the various stages of mature capsid assembly.

Viral DNA is synthesized via a rolling-circle mechanism (reviewed in reference 29) in a central nuclear compartment that has been defined as the presence of the single-stranded viral DNA-binding protein (19). The newly replicated DNA is cleaved into unit-length genomic form and packaged into the preformed, immature capsids, which then acquire an envelope derived from the inner nuclear membrane upon exit from the nucleus. In this report we compare the distribution of capsid proteins with the HSV-1 DNA polymerase accessory protein encoded by the U_L42 gene, as well as the U_L15 gene product, which has been associated with cleavage and packaging of viral DNA on the basis of analyses of a spontaneously arising mutant virus carrying a temperature-sensitive (*ts*) mutation in the U_L15 gene. At the nonpermissive temperature, cells infected with this mutant accumulate uncleaved concatemeric viral DNA, and packaging into preformed capsids does not ensue (1, 26). Several other viruses which contain *ts* mutations in genes whose products are either capsid proteins or proteins known to be required for viral DNA synthesis accumulate uncleaved viral DNA (36).

Previously, we observed that ICP35 protein was distributed in discrete patches at the periphery of the nucleus late in infection (44). We asked whether these discrete patches also contained proteins associated with mature capsids, thereby suggesting that these sites could be sites of capsid assembly rather than represent a concentration of soluble proteins. We therefore compared the distribution of ICP35 with that of ICP5 and U_L38 (VP19c). Relevant to this are the following points. (i) The U_L26 gene encodes a protease precursor, Pra (19). A transcriptional unit designated $U_L26.5$, 3′ coterminal with the U_L26 gene, yields ICP35cd (18). Both the protease precursor Pra and the ICP35cd protein are cleaved by the protease to yield several sets of products. ICP35cd is cleaved to yield ICP35ef and a small carboxy-terminal peptide (19). Pra is cleaved to yield the mature protease, a polypeptide designated ICP35ab, and the same small carboxy-terminal peptide as that cleaved from ICP35cd (10, 12, 20). ICP35ab, ICP35ef, and Prn (the amino-terminal cleavage product of Pra) form the scaffolding of the capsid. Upon packaging of viral DNA, only Prn (VP21) remains in the capsid (14, 25). Monoclonal antibody H725 reacts with an epitope present in Pra, ICP35ab, ICP35cd, and ICP35ef but not in Prn (3, 4, 18). This antibody therefore

* Corresponding author. Present address: The Marjorie B. Kovler Viral Oncology Laboratories, University of Chicago, 910 E. 58th St., Chicago, IL 60637. Phone: (312) 702-1693. Fax: (312) 702-1631.

identifies accumulations of U_L26 gene products either dispersed or contained in [illegible] ids. (ii) In contrast, ICP5, the major capsid protein, [illegible] ment of both immature and mature capsids, and [illegible] 5 identifies the accumulation of soluble prot[illegible] s the accumulation of both immature and mature c[illegible]. (iii) The U_L38 gene encodes a capsid protein [illegible] d VP19c (5, 16, 21). This protein has been previ[illegible]y s[illegible]n to bind viral DNA in a nonspecific fashion and [illegible]y [illegible] a role in anchoring the DNA to the capsid (5, [illegible]5).

The site of acquisi[illegible]n of the tegument is uncertain. The observations that our [illegible] from nuclei do not contain tegument pr[illegible]s (7[illegible] [illegible]ated patches of membrane form at the s[illegible] [illegible] at the inner nuclear membrane have led t[illegible] [illegible] that tegument proteins are part of the envelope [illegible]ed at the nuclear membrane (35). The loc[illegible]ion [illegible]ment proteins has been investigated in this stud[illegible]e. [illegible]ct of the U_S11 gene is an abundant tegument p[illegible] [illegible] binds RNA in a sequence- and conformation-spe[illegible] [illegible] binds to ribosomes, and also localizes in the n[illegible] (3[illegible]–35). The alpha-*trans*-inducing factor [illegible] known as virion protein 16 (42), transactivates t[illegible] [illegible]t the α genes—the first set of genes ex[illegible]s[illegible] [illegible] infected cells (28, 36). αTIF is also [illegible]sse[illegible] [illegible]cia. Virions produced in cells in[illegible]d with [illegible] generated by substitution of cyste[illegible]e with [illegible]ol glycines fail to mature at the nor[illegible]ve [illegible] (2[illegible]).

In this r[illegible]port we [illegible]g. (i) Viral proteins involved in the [illegible]C[illegible]esis [illegible] are dispersed throughout a large po[illegible]ti[illegible]e [illegible] proteins associated with immature ca[illegible]s[illegible] [illegible]. and U_L38 (VP19c), aggregate in d[illegible] [illegible] nuclear structures frequently loc[illegible]d [illegible]r, diffuse nuclear domains defined [illegible] DNA synthesis. (iii) The U_L15 [illegible] was required for packaging of viral D[illegible]A [illegible] itself a capsid protein, localizes w[illegible] [illegible] DNA synthesis and surprisingly [illegible] U_L[illegible] capsid proteins. We conclude [illegible] that late in infection at the time of h[illegible]g[illegible]me [illegible] virions, components of the capsid and s[illegible]eg[illegible] assemble in discrete nuclear struc[illegible]s[illegible] [illegible] space occupied by proteins involved i[illegible] D[illegible] have designated these structures [illegible] y include proteins of both mature and in[illegible] also some tegument proteins.

MA[illegible]ODS

Cells and virus[illegible] [illegible] the prototype strain used in our laborato[illegible] [illegible]. The U_S8 gene which encodes glycoprotein [illegible] was [illegible] ace studies to avoid nonspecific immu[illegible] [illegible]s[illegible] [illegible] immunoglobulin G (IgG) to the Fc receptor [illegible] [illegible] V-1(F)R2063 carries two *ts* mutations in [illegible] in the *ul* gene is repaired (27). HSV-1[illegible] [illegible] U_L15 gene (26). At the nonpermissi[illegible] [illegible] remains in concatemers and is not [illegible] 2%. All experiments were done in Vero c[illegible] [illegible] MEM medium supplemented with 5% newborn [illegible]

Antibodies. [illegible] U_L10 (glycoprotein M) (2), U_L15 (1), and [illegible] described previously. Rabbit polyclonal [illegible] generated by immunization of rabbits wi[illegible] [illegible] below. The preparation of the monoclo[illegible] [illegible] described elsewhere (39). Monoclonal [illegible]5 and ICP35 were obtained from Good[illegible] [illegible] monoclonal antibody LP1, which reacts wi[illegible] [illegible] of A. Minson. The monoclonal antibody [illegible] encoded by the U_L42 gene was the kind gift of Dan Tenney, Bristol-Meyers Squibb. The goat anti-rabbit fluorescein isothiocyanate (FITC)-conjugated antibody was purchased from Sigma Chemical Co., St. Louis, Mo. The goat anti-mouse Texas red-conjugated antibody was purchased from Molecular Probes, Inc., Eugene, Oreg.

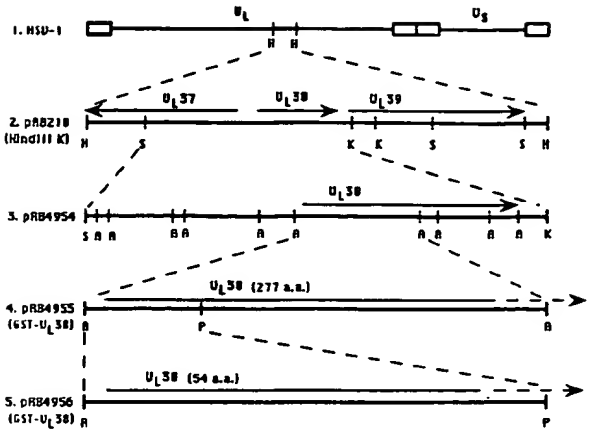


FIG. 1. Schematic representation of the sequence arrangements of HSV-1 (F) DNA and of the plasmids used to construct the GST fusion with the U_L38 ORF. Line 1, sequence arrangement of HSV-1(F) DNA. Open boxes, internal repeat sequences flanking the unique long (U_L) and unique short (U_S) regions; line 2, sequence arrangement of pRB210 containing the U_L37, U_L38, and U_L39 ORFs located within the *Hin*dIII K fragment of HSV-1(F) DNA; line 3, sequence arrangement of pRB4954 containing the *Sal*I-*Kpn*I fragment from pRB210 cloned into the pGEM3Z vector and encompassing the entire U_L38 ORF and a portion of the U_L37 ORF; line 4, 864-bp *Ava*I fragment from pRB4954 encoding the N-terminal 277 amino acids of U_L38 cloned into the *Ava*I site of vector pGEX-2T, creating pRB4955; line 5, pRB4955 digested with pPuMI and *Eco*RI (site present in vector polylinker), with the ends made blunt and religated, reducing the U_L38 portion of the GST-U_L38 fusion protein to 54 amino acids (a.a.). A, *Ava*I; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; P, pPuMI; S, *Sal*I.

Preparation of anti-U_L38 polyclonal antiserum. Figure 1 illustrates the construction of plasmid pRB4956 encoding the U_L38–glutathione-*S*-transferase (U_L38-GST) fusion protein. Line 1 shows the sequence arrangement of the HSV-1 genome, and line 2 shows the sequence arrangement of HSV-1(F) *Hin*dIII-K cloned into the *Hin*dIII site of pBR322. The resultant plasmid, designated pRB210, contains the U_L37, U_L38, and U_L39 open reading frames (ORFs). The 3.1-kbp *Sal*I-*Kpn*I fragment of pRB210 was ligated into the *Sal*I and *Kpn*I sites in pGEM3Z (Promega, Madison, Wis.) to yield pRB4954 (Figure 1, line 3). The 864-bp *Ava*I fragment of pRB4954 encoding the amino-terminal 277 amino acids of the U_L38 ORF was ligated into the *Ava*I site of pGex2T (Pharmacia). The resultant plasmid, designated pRB4955 (Fig. 1, line 4), was predicted to encode the bacterial GST in frame with the amino-terminal region of U_L38. pRB4955 was digested with pPuMI and *Eco*RI (the *Eco*RI site is in the vector polylinker), and the ends were made blunt with mung bean nuclease (New England Biolabs, Beverly, Mass.) and religated. The resultant plasmid, designated pRB4956, contains the sequence encoding the amino-terminal 54 amino acids of U_L38 fused to the GST gene (Fig. 1, line 5). DNA encoding the junction between GST and U_L38 was sequenced to verify that the two open ORFs were maintained (data not shown). Production of the fusion protein was induced by the addition of IPTG (isopropyl-β-D-thiogalactopyranoside) to the medium with *Escherichia coli* BL21 cells transformed with pRB4956, followed by affinity purification with glutathione cross-linked to agarose beads (Sigma), and checked for purity by separation on polyacrylamide gels followed by staining with Coomassie brilliant blue. New Zealand White rabbits were immunized with the affinity-purified protein as previously described (2). Samples of preimmune and immune serum were collected and analyzed for reactivity to U_L38 protein by immunoblotting.

Polyacrylamide gel electrophoresis and immunoblotting. Infected cell lysates were separated in denaturing gels consisting of 10% polyacrylamide and 0.2% sodium dodecyl sulfate, and the proteins were electrically transferred to nitrocellulose sheets. The sheets were soaked at room temperature for 1 h in phosphate-buffered saline (PBS) containing 5% skim milk (Carnation) and then reacted at room temperature for 1 h with the U_L38 polyclonal antiserum (1:1,000 dilution) or with preimmune serum (1:500 dilution) in PBS containing 1% bovine serum albumin (BSA). The blots were washed three times for 5 min in PBS containing 5% milk and then reacted at room temperature for 1 h with a 1:3,000 dilution of goat anti-rabbit antibody conjugated to alkaline phosphatase. The blots were washed for 10 min in PBS containing 5% milk and then four times for 10 min in PBS; they were then developed by using reagents and protocols

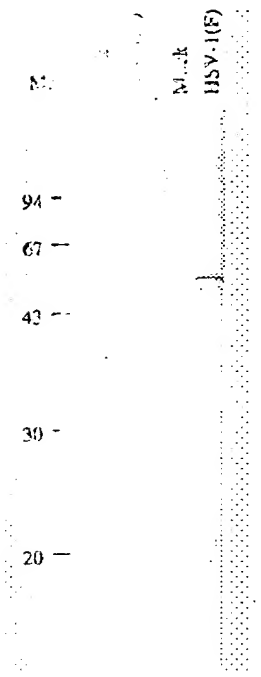


FIG. 2. Photograph of ir... horetically separated lysates of HSV-1(F)- or mock-infect... 2 h postinfection. The blots were probed with rabb... U_L38-GST fusion protein (right panel) or with preir... . Molecular weights are in thousands.

supplied in a kit from the ... Bio-Rad Laboratories, Richmond, Calif.).

Immunofluorescence. Ap... Vero cells were seeded onto glass slides (Cell-line Inc., N... attach overnight, and then exposed to 10 PFU of virus ... h of incubation at temperatures stated in Results, the... old methanol, air dried, and reacted with PBS or PBS ... 1% BSA and 20% normal human plasma ... nonspecific reactivity of the antibodies. The cells were ... temperature with primary antibody diluted in PBS ... alone or along with 10% normal human plasma ... reacted for 1 h with the appropriate secondary ant... and mounted in 95% glycerol in PBS containing 1 mg of ... ml to reduce fading of the FITC signal. The slides were ... confocal fluorescence microscope; digitized images ... stained cells were acquired with software produced ... and printed by a CP210 Codonics printer. S... by excitation with an argon/krypton laser ... ed). Double-stained images were obtained ... corochromes filtered by 515–540-nm ... ss (Texas red) filters and subsequent ...

Specificity of the ... clonal antiserum. We generated a rabbit po... to U_L38 protein for use in colocalization ... monoclonal antibodies to ICP5 ... oduced following inoculation of ... n protein specifically reacted ... 53 kDa present in HSV-1(F) ... t panel). The antibody did ... ent in mock-infected cell lysates ... e preimmune sera react with any proteins pres... fected or infected cell lysates (left panel).

Redistribution of IC... during HSV-1 infection. The localization of ... ciated with immature (ICP35) and mature (ICP5) capsids (33, 36) was examined at early (6 h) and late (16 h) times postinfection. At 6 h postinfection, both ICP5 (Fig. 3A) and ICP35 (Fig. 3D) proteins are diffusely distributed in an irregular pattern throughout the nucleus. In some cells, some aggregation of capsid proteins can be seen (Fig. 3A, compare cells labeled a and b), which progresses to form discrete, brightly fluorescent structures by 16 h postinfection (Fig. 3B [ICP5], cell labeled a; Fig. 3C [ICP5]; Fig. 3E and F [ICP35]). The localization of the tegument protein, αTIF (Fig. 3G to I), is discussed below.

Colocalization of U_L38, ICP35a-f, and ICP5. As previously reported (44), the $U_L43.5$ protein colocalizes with ICP35 in dense, strongly fluorescent nuclear structures (Fig. 4a to c). In this series of experiments, we asked whether these structures contained exclusively proteins associated with immature capsids by comparing the localization of three proteins associated with mature (ICP5 and U_L38) and immature (ICP5, U_L38, and ICP35) capsids. The U_L38 protein was detected with FITC-conjugated antibody against rabbit IgG. ICP5 and ICP35 were detected with Texas red-conjugated antibody against mouse IgG. The results of these studies show that in some cells, ICP35 and ICP5 (Fig. 4g, j, and m) aggregated in both diffuse and highly dense, strongly fluorescent nuclear structures. Colocalization of these proteins with the U_L38 protein is visualized by yellow fluorescence (Fig. 4, right-hand column). U_L38 protein colocalized both with ICP35 and ICP5 in the dense, strongly fluorescent nuclear structures (Fig. 4k, l, n, o, q, and r) and with the diffuse, less strongly fluorescent ICP5 (Fig. 4n and o) but not with the diffusely distributed ICP35 (Fig. 4h and i). In Fig. 4, we have illustrated the dominant pattern consisting of a small number of dense structures (three to eight) prevalent in the infected-cell nuclei late in infection. In some cells, particularly in those in which infection was retarded, the nuclei contained a large number of relatively smaller dense, fluorescent structures. This is illustrated in Fig. 5b and c. To define the localization of these dense structures more precisely, infected cells were reacted simultaneously with rabbit polyclonal antibody to glycoprotein M and with monoclonal antibody to ICP35. As illustrated in Fig. 6g to i, the dense, strongly fluorescent nuclear structures containing ICP35 were separated from the uniform shell formed by glycoprotein M in the nuclear membrane.

Tegument proteins partially overlap the dense, strongly fluorescent nuclear structures containing capsid proteins. We have examined the localization of two tegument proteins, U_S11 and αTIF. U_S11 protein accumulated in diffuse nuclear regions at or near nuclear membranes in most of the infected cells, in addition to nucleoli as previously reported (39). Even in cells in which U_L38 protein is largely in the dense, strongly fluorescent nuclear structures, the U_S11 protein overlapped only in part with these structures containing capsid proteins (Fig. 4d to f).

The localization of αTIF was examined under five conditions, i.e., in cells infected with HSV-1(F) or with viruses carrying *ts* mutations in U_L15 [HSV-1(mP)*ts*66-4] or in αTIF (R2603) and maintained at either permissive or nonpermissive temperatures. The results of these studies were as follows. (i) Abrogation of DNA cleavage and packaging did not affect the appearance of any of the compartments defined by the presence of capsid proteins or proteins associated with DNA synthesis in cells infected with this mutant virus at the nonpermissive temperature. (ii) As a general rule, αTIF was an abundant nuclear protein that was present either diffusely throughout the nucleus (Fig. 5a) with accumulation of protein in nuclear membranes (Fig. 3G and H) or aggregated and only partially filling the nucleus (Fig. 3I). The compartment occupied by αTIF was considerably more diffuse and only partially colocal-

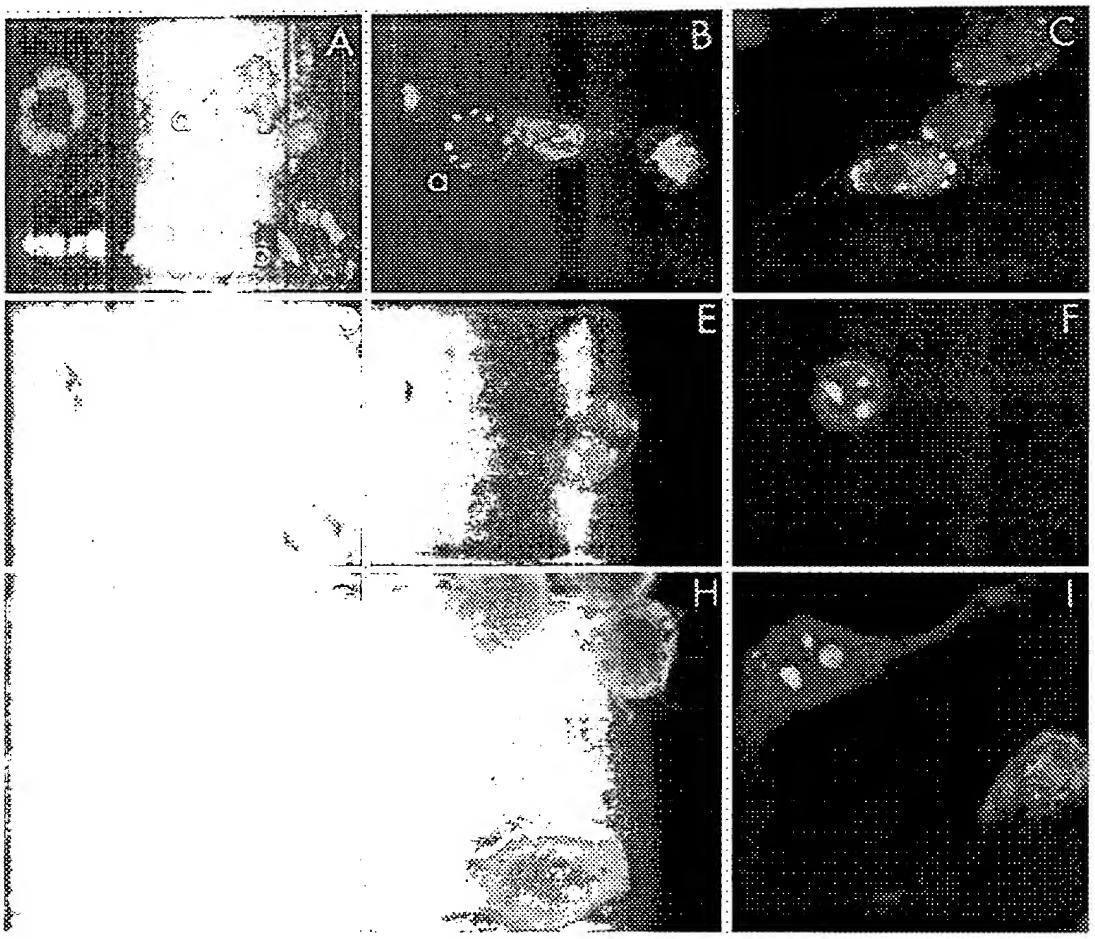


FIG. 3. (... or R7032(gE-)-infected (A to F) Vero cells fixed at 6 h (A and D) or 16 h (B, C, and E to I) postinfectio... D to F), or to a HF protein (G to I) and to anti-mouse IgG conjugated to Texas red. The images were captu... and printed by a Codonics CP210 printer.

ized with... ...ther capsid proteins ... ...perature for the ... ...es of the infected ... ...localized within th... ...e margins of the nu... ...es (Fig. 5d). In t... ...ap in the localizati... ...clear struc-tures cor... ...s indicated below), ot... ...no change in distributi... ...ells infected with R2... ...ve tempera-ture (dat...

The lo... ...hat of pro-**teins ass...** ...roteins as-sociated with DNA replication accumulate in a specific nuclear compartment that is defined by the presence of the viral single-stranded-DNA-binding protein ICP8 (9). We examined the distribution of another protein associated with viral DNA synthesis, the polymerase accessory protein encoded by U_L42. Our observations, supported by other colocalization studies described below, are that the viral DNA replication compartment abuts but does not overlap the dense nuclear fluorescent structures containing capsid proteins (Fig. 6a to c). Studies with monoclonal antibody to ICP8 yielded similar results (not shown).

Colocalization of U_L15 protein with proteins involved in viral DNA synthesis. Earlier studies have shown that U_L15 protein is redistributed from the cytoplasm to the nucleus between 6 and 12 h after infection (1). In this study, the

FIG. 4. (... d ... cells fixed and stained 16 h postinfection and double labeled with combinations of antibodies to viral proteins an... ... fluorescence) or an... rabbit IgG conjugated to FITC (green fluorescence). Single-color images were captured s... ...s; the two colors were then captured simultaneously and are shown in the right column. The yellow color visua... ...s colocalization of red and green fluorescence. (a to c) ICP35 (red) and $U_L43.5$ (green); (d to f) U_S11 (red) and U_L... ...j to l, ICP35 (red) and U_L38 (green); (m to o) ICP5 (red) and U_L38 (green); (p to r) ICP5 (red) and U_L... ...d by Zeiss with the instrument and printed by a Codonics CP210 printer. The abundance of $U_L43.5$ protein is l... ...e of $U_L43.5$ was digitally enhanced to match that of ICP35.

FIG. 5. [illegible] in the U_L15 ger[illegible] ed Vero cells incubated at the permissive (a to c and g to l) or nonpermissive (d to f) ten[illegible] [illegible]uble labeled [illegible]binations of antibodies to viral proteins and with the appropriate secondary antibody co[illegible] [illegible]en fluor[illegible] Single-color images are represented in the left and middle columns; the right column rep[illegible] [illegible]olor visu[illegible] [illegible]e overlaid images (right column) represents colocalization of red and green fluorescen[illegible] [illegible]ody; (g t[illegible] [illegible]o l) ICP5 (red) and U_L15 (green), respectively. The images were captured with softw[illegible] [illegible]C Jonics C[illegible] [illegible]ater. The fluorescence signal obtained for the individual proteins shown in this figure [illegible]

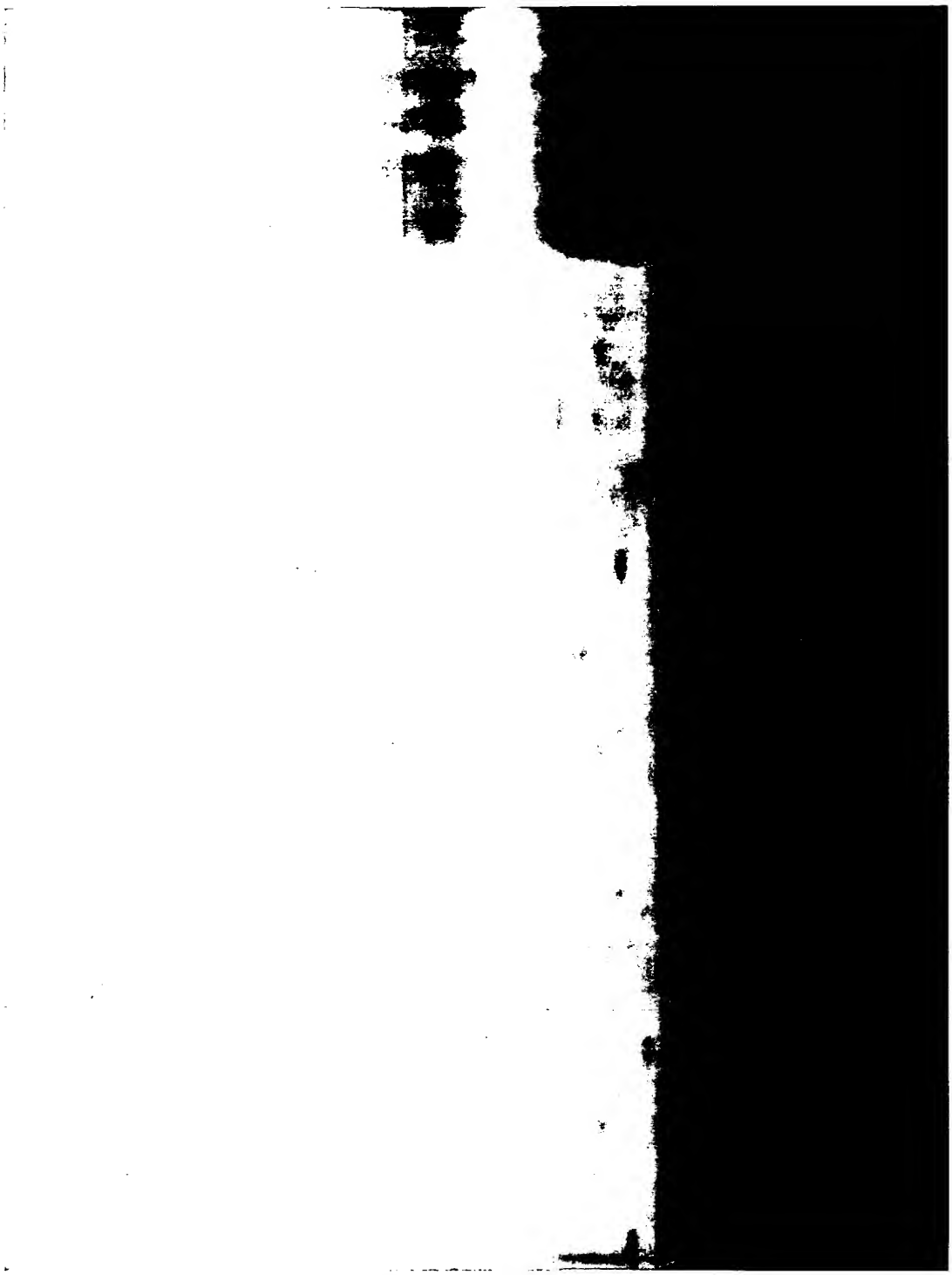


FIG. 6. (... is fixed at 16 ... ection and double labeled with combinations of antibodies to viral proteins and with th... ...ed (red ...ne) or FITC (green fluorescence). Single-color images are represented in the left an... ...quisi... ...colors. The yellow color in the overlaid images (right column) represents colocalizat... ... ICP35 (red) and U_L15 (green); (g to i) ICP35 (red) and gM (green); (j to l) U_L42 ... ...ware pr... ...eiss with the instrument and printed by a Codonics CP210 printer. The fluorescen... ...gure was ...nd did not require digital enhancement.

localization of the
1(F) and
tained at
investigat
protein in
temperatu
amounts o
cupied the
1(mP)*ts*6
bution of
5h). In ce
peripheral
the nucle
U_L15 pr
localizati
taining ca
lapping t
Fig. 5g to
observatio
teins asso

In a recen
newly dis
sid prote
we have
A central
was uniqu
compart
viral mo
ing of D
that a fu
antibody to
encoded
three cap
proteins
synthesis
U_L42 ge
the virion
gene (2).
occurred
packagin
with mut
(Fig. 5) or
nonperm

We ha
distinct
These c
volved i
recently
aggregate
ment.di
of our st
capsid c
ICP35, an
which occ
were loc
nuclear
partme
(ii) U_L1
age and
cessory
dense b
dominan

fected with HSV-
*ts*6-4 and main-
temperatures was
ribution of U_L15
dependent of the
n, relatively large
-U_L15 antibody oc-
F), Fig. 6e; HSV-
pattern, the distri-
e diffuse (e.g., Fig.
ve to the discrete
e amounts, filling
's location of the
nage is that of
structures con-
partially over-
$_L15$ protein (e.g.,
and unexpected
zed with the pro-
, U_L42 [Fig. 6j to
l]).

roduct of $U_L43.5$, a
es with the cap-
rtments which
structures (44).
er this association
reflected a nuclear
s associated with
ssembly, packag-
te the possibility
sts, we prepared
, VP19C (15, 42),
the localization of
, two tegument
d in viral DNA
the product of the
and a component of
product of the U_L10
) a fundamentalization
ctions, i.e., DNA
ning cells infected
either the U_L15
and maintained at

eins in which three
ear to occur.
proteins in-
ins. We have
NA synthesis, ICP4
form a compart-
adient features
follows. (i) The
sence of ICP5,
ear structures,
These structures
us separate from
lap the com-
NA synthesis.
viral DNA cleav-
merase ac-
not with the
(iii) The pre-
proteins αTIF and

U_S11 was in asymmetrically arranged masses near the periphery of the nucleus, adjacent to or partially overlapping with the dense nuclear structures containing capsid proteins, although the distribution of the two proteins varied somewhat from cell to cell. We did not observe significant differences in the compartmentalization of the proteins studied here in cells in which either DNA packaging was abrogated (U_L15 *ts* mutant) or virions did not mature (αTIF mutant).

The hypothesis that the dense, strongly fluorescent structures contain both protein forming immature capsids and the immature capsids themselves, in addition to any mature capsids which may also be present, is supported by two observations: (i) the presence of large amounts of ICP35 protein within these structures and (ii) the formation of these structures in infected cells under conditions in which immature capsids (lacking viral DNA) are made, viral DNA accumulates in the form of concatemers, but the DNA is not cleaved from concatemers and packaged into the preformed capsids (1, 26). At the time points investigated in this report, infectious virus accumulates exponentially, and we assume that this represents a time of maximum assembly of capsids. "Dense fluorescent nuclear structures," albeit descriptive, is an unsatisfactory term for these structures, and we propose to designate them assemblons, i.e., bodies in which proteins involved in capsid assembly aggregate and most probably assemble.

The arrangement of assemblons in the periphery of the nucleus surrounding and abutting the compartment containing proteins involved in viral DNA synthesis suggests the existence of a machinery which feeds upon (i.e., captures, packages, and cleaves) viral DNA generated within the DNA synthesis compartment. The distribution of U_L15 protein raises some interesting issues. U_L15 bears partial homology to the bacteriophage terminase involved in cleavage and packaging of phage DNA (7, 29, 32), and analyses of a *ts* mutant virus revealed the requirement for functional U_L15 protein for cleavage and packaging of HSV-1 DNA into capsids (1, 26). Because it is generally thought that cleavage of newly synthesized viral DNA occurs concurrently with its packaging (reviewed in reference 36), it could have been predicted that U_L15 would colocalize with immature capsids or with structures associated with capsid assembly. Thus, the colocalization of U_L15 protein with proteins involved in viral DNA synthesis and, conversely, the virtually complete exclusion of U_L15 protein from the assemblons were unexpected. Our results suggest that U_L15 becomes associated with DNA-protein complexes prior to packaging and that it may even be an accessory component of the machinery which synthesizes viral DNA.

In contrast, and with few exceptions, the two tegument proteins investigated in this study (e.g., Fig. 5d to f) were localized closer to nuclear membranes than were the assemblons. We have not resolved the question of whether tegument proteins bind to capsids during capsid assembly or at the time of envelopment. Both U_S11 and αTIF are abundant virion proteins; however, our results suggest that the assemblons do not contain large amounts of U_S11 or αTIF as might be expected if tegument proteins bound to capsids within these structures.

The conclusions drawn from these studies have generated a number of questions. In principle, we could have expected that assembly of capsids in the nucleus would occur at random but that the capsids would move radially to the nuclear membrane. The fact that the capsid proteins accumulate in assemblons suggests that assembly is facilitated by and is dependent on protein concentration. This raises the question of the mechanisms which result both in the coalescence of the capsid proteins into assemblons and in the segregation of proteins involved in DNA synthesis and those forming the tegument.

An addi[illegible]nal, e[illegible] ·stion concerns the re-
lationship [illegible] mpartments and the
electron-de[illegible]se b[illegible] apparent upon elec-
tron-micros[illegible] ons of infected cells
(31, 40). Pr[illegible] components of these
bodies will r[illegible] ethods which enable
penetration of an[illegible] out distortion of the
nuclear architectur[illegible]

A 28

We thank Lindsay S[illegible] nical assistance and A.
Minson a[illegible] and anti-U_L42 anti-
bodies, res[illegible]
These s[illegible] National Cancer Insti-
tute (CA4[illegible] Allergy and Infectious
Diseases (A[illegible] ice.

1. Baines, J. D., A. P. [illegible] Roizman. 1994. The herpes simplex vi[illegible] and is required for cleavage of genomic viral [illegible]
2. Baines, J. D., and [illegible] ene of herpes simplex virus 1 encodes [illegible] present in the virion and in the [illegible]. 67:1441–1452.
3. Brau[illegible] n. 1983. Application of denat[illegible] ized lysates of herpes simple[illegible] noclonal antibodies and for stu[illegible] ol. 45:103–112.
4. Braun, [illegible] haracterization of post-transl[illegible] proteins binding to the surfaces o[illegible] Virol. 49:142–153.
5. Braun, D. [illegible] 4. Identification and genetic manipi[illegible] that binds DNA. J. Virol. 50:6[illegible]
6. Costa, R. [illegible] d E. Wagner. 1984. Direct demonst[illegible] pes simplex virus type 1 mRNA [illegible] encodes the major capsid protei[illegible]
7. Dav[illegible] e of herpesvirus. Virology [illegible]
8. Davison [illegible] 92. Identification of genes encod[illegible] pes simplex virus type 1. J. [illegible]
9. Delb[illegible] g nuclear architecture defines [illegible] NA replication structures. J. Virol. [illegible]
10. Deckman, [illegible] 1992. Herpes simplex virus type 1 p[illegible] is autoprocessing and specific[illegible] Virol. 66:7362–7367.
11. Diaz, J.-[illegible] Albeiter, L. Denoroy, and J.-J. M[illegible] U_S11 gene product is a phosph[illegible] ily associated with both ribos[illegible]
12. Dil[illegible] , McCann III, F. Liu, B. Roizm[illegible] ntification of the herpes [illegible] Acad. Sci. USA 68:2048–2051.
13. Fjerdi[illegible] s. Characterization of herp[illegible] ts on social behavior of infected [illegible]
14. Gibson, [illegible] ecified by herpes simplex virus. V[illegible] ltiple capsid forms of subtypes [illegible]
15. Gibson, [illegible] ified by herpes simplex virus. X[illegible] B-capsid and virion proteins [illegible]
16. Heine[illegible] , Roizman. 1974. Proteins specif[illegible] n polypeptides of type 1 strains [illegible]
17. Lee, an[illegible] rection and co-localization [illegible] of ICP4 with EAP, a nuc[illegible] l. USA 93:4583–4587.
18. Liu, F., [illegible] nscriptional unit, and coclin[illegible] proteins are contained with[illegible] ac. Virol. 65:206–212.
19. Liu, F., [illegible] virus 1 gene encoding a prote[illegible] ne encoding the more abun[illegible]
20. Liu, F., [illegible] of the protease and other products of amino-terminus-proximal cleavage of the herpes simplex virus 1 U_L26 protein. J. Virol. 67:1300–1309.
21. McGeoch, D. J., M. A. Dalrymple, A. J. Davison, A. Dolan, M. C. Frame, D. McNabb, L. J. Perry, J. E. Scott, and P. Taylor. 1988. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. 69:1531–1574.
22. McNabb, D. S., and R. J. Courtney. 1992. Identification and characterization of the herpes simplex virus type 1 virion protein encoded by the U_L35 open reading frame. J. Virol. 66:2653–2663.
23. Meignier, B., R. Longnecker, P. Mavromara-Nazos, A. Sears, and B. Roizman. 1988. Virulence of and establishment of latency by engineered deletion mutants of herpes simplex virus 1. Virology 162:251–254.
24. Pellet, P. E., J. L. C. McKnight, F. J. Jenkins, and B. Roizman. 1985. Nucleotide sequence and predicted amino acid sequence of a protein encoded in a small herpes simplex virus DNA fragment capable of trans-inducing α genes. Proc. Natl. Acad. Sci. USA 82:5870–5874.
25. Person, S., S. Laquerre, P. Desai, and J. Hempel. 1993. Herpes simplex virus type 1 capsid protein, VP21 originates within the U_L26 open reading frame. J. Gen. Virol. 74:2269–2273.
26. Poon, A. P. W., and B. Roizman. 1993. Characterization of a temperature-sensitive mutant of the U_L15 open reading frame of herpes simplex virus 1. J. Virol. 67:4497–4503.
27. Poon, A. P. W., and B. Roizman. 1995. The phenotype in vitro and in infected cells of herpes simplex virus 1 α *trans*-inducing factor (VP16) carrying temperature-sensitive mutations introduced by substitution of cysteines. J. Virol. 69:7658–7667.
28. Post, L. E., S. Mackem, and B. Roizman. 1981. Regulation of α genes of herpes simplex virus: expression of chimeric genes produced by fusion of thymidine kinase with α gene promoters. Cell 24:555–565.
29. Powell, D., J. Franklin, F. Arisaka, and G. Mosig. 1990. Bacteriophage T4 DNA packaging genes 16 and 17. Nucleic Acids Res. 18:4005.
30. Preston, V. G., J. A. V. Coates, and F. J. Rixon. 1983. Identification and characterization of a herpes simplex virus gene product required for encapsidation of virus DNA. J. Virol. 45:1056–1064.
31. Puvion-Dutilleul, F., J. Pedron, M. L Althier, and P. Sheldrick. 1982. Ultrastructural studies on the nucleus of herpes simplex virus type 1-infected cells. Biol. Cell 44:249–260.
32. Rao, V. B., and L. W. Black. 1988. Cloning, overexpression, and purification of the terminase proteins gp16 and gp17 of bacteriophage T4. Construction of a defined in vitro DNA packaging system using purified terminase proteins. J. Mol. Biol. 200:475–488.
33. Rixon, F. J., A. M. Cross, C. Addison, and V. G. Preston. 1988. The products of herpes simplex virus type 1 gene U_L26 which are involved in DNA packaging are strongly associated with empty but not with full capsids. J. Gen. Virol. 69:2879–2891.
34. Rixon, F. J., M. D. Davison, and A. J. Davison. 1990. Identification of the genes encoding two capsid proteins of herpes simplex virus type 1 by direct amino acid sequencing. J. Gen. Virol. 71:1211–1214.
35. Roizman, B., and D. Furlong. 1974. The replication of herpesviruses, p. 229–403. *In* H. Fraenkel-Conrat and R. R. Wagner (ed.), Comprehensive virology, vol. 3. Plenum Press, New York.
36. Roizman, B., and A. E. Sears. 1996. The replication of herpes simplex viruses, p. 2231–2295. *In* B. N. Fields, D. M. Knipe, P. Howley, R. M. Chanock, M. S. Hirsch, J. L. Melnick, T. P. Monath, and B. Roizman (ed.), Virology, 3rd ed. Raven Press, New York.
37. Roller, R. J., and B. Roizman. 1990. The herpes simplex virus U_S11 open reading frame encodes a sequence-specific RNA-binding protein. J. Virol. 64:3463–3470.
38. Roller, R. J., and B. Roizman. 1991. Herpes simplex virus 1 RNA binding protein U_S11 negatively regulates the accumulation of a truncated viral mRNA. J. Virol. 65:5873–5879.
39. Roller, R. J., and B. Roizman. 1992. The herpes simplex virus 1 RNA binding proteins U_S11 is a virion component and associates with ribosomal 60S subunits. J. Virol. 66:3624–3632.
40. Schwartz, J., and B. Roizman. 1969. Similarities and differences in the development of laboratory strains and freshly isolated strains of herpes simplex virus in HEp-2 cells: electron microscopy. J. Virol. 4:879–889.
41. Sherman, G., and S. L. Bachenheimer. 1987. DNA processing in temperature-sensitive porphogenic mutants of HSV-1. Virology 158:427–430.
42. Spear, P. G., and B. Roizman. 1972. Proteins specified by herpes simplex virus. V. Purification and structural proteins of the herpesvirion. J. Virol. 9:143–159.
43. Spector, D., F. C. Purves, R. W. King, and B. Roizman. 1993. Regulation of β and γ gene expression in cells infected with herpes simplex viruses, p. 25–42. *In* L. Carrasco et al. (ed.), Regulation of gene expression in animal viruses. Plenum Press, New York.
44. Ward, P. L., D. Barker, and B. Roizman. 1996. A novel herpes simplex virus 1 gene U_L43.5 maps antisense to the U_L43 gene and encodes a protein which colocalizes in nuclear structures with capsid proteins. J. Virol. 70:2684–2690.
45. Zweig, M., C. J. Heilman, and B. Hampar. 1979. Identification of disulfide-linked protein complexes in the nucleocapsids of herpes simplex virus type 2. Virology 94:442–450.


a
ICP35
b
UL43.5
c
e
f
h
i
l
o
r

a
b
c
e
f
i
l